Neural basis of concurrent deliberation toward a choice and degree of confidence

Miguel Vivar-Lazo[1] & Christopher R. Fetsch[1,2]

[1]Zanvyl Krieger Mind/Brain Institute, [2]Solomon H. Snyder Department of Neuroscience,

Johns Hopkins University, Baltimore, MD, USA; correspondence: cfetsch@jhu.edu

SUMMARY

Decision confidence plays a key role in flexible behavior, but exactly how and when it arises in the brain remains unclear. Theoretical accounts suggest that confidence can be inferred from the same evidence accumulation process that governs choice and response time (RT), implying that a provisional confidence assessment could be updated in parallel with decision formation. We tested this using a novel RT task in nonhuman primates that measures choice and confidence with a single eye movement on every trial. Monkey behavior was well fit by a 2D bounded accumulator model instantiating parallel processing of evidence, rejecting a serial model in which choice is resolved first followed by post-decision accumulation for confidence. Neural activity in area LIP reflected concurrent accumulation, exhibiting within-trial dynamics consistent with parallel updating at near zero time lag, and significant covariation in choice and confidence signals across the population. The results demonstrate that monkeys concurrently process a single stream of evidence to arrive at a choice and level of confidence, and illuminate a candidate neural mechanism for this ability.

INTRODUCTION

Decisions are accompanied by a sense of confidence, defined as the degree of belief that the decision is correct. Confidence facilitates learning in the absence of explicit feedback (Daniel & Pollmann, 2012; Guggenmos et al., 2016) and guides decisions that are part of a sequence or hierarchy (Sarafyazd & Jazayeri, 2019; Purcell & Kiani, 2016; van den Berg et al., 2016b). When feedback does occur, confidence informs whether the outcome is surprising (i.e., a high confidence error), driving an increase in learning rate (Rescorla & Wagner, 1972; Sutton & Barto, 1992).

Indeed, it can be shown that the optimal weights for converting sensory neuron activity to decision evidence can only be obtained with a learning rule that is proportional to confidence (Drugowitsch et al., 2019). Even when subjects are well-trained and task contingencies are stable, choice biases continuously evolve in response to previous outcomes in a difficulty-dependent manner, suggesting a role for confidence (Lak et al., 2020). These facts imply that confidence is of broad importance for understanding sensorimotor behavior and learning, yet we still lack a mechanistic understanding of its neuronal underpinnings. Here we focus on the timing of confidence computations in relation to the primary decision (Xue et al., 2023; Herregods et al., 2023; Moran et al., 2015; Pleskac & Busemeyer, 2010), with the goal of uncovering the latent dynamics of choice and confidence signals in the brain (Murphy et al., 2015; Gherman & Philiastides, 2015; Pereira et al., 2020; Balsdon et al., 2021; Wang et al., 2023; Dou et al., 2024; Fan et al., 2024). Beyond informing theories of decision making and learning, this question has implications for the computational and neural basis of metacognition.

Most existing models fall into one of two main alternatives for the dynamics of choice and confidence formation. The *serial* hypothesis states that the two computations make use of temporally nonoverlapping streams of evidence (Herregods et al., 2023; Moran et al., 2015; Pleskac & Busemeyer, 2010). In contrast, the *parallel* hypothesis proposes simultaneous initiation and temporally overlapping (concurrent) computation of choice and confidence (Dotan et al., 2018; van den Berg et al., 2016a; Xue et al., 2023) — though this need not exclude additional post-decision processing (Desender et al., 2021; Maniscalco et al., 2021; see Discussion). Both hypotheses have received support from behavioral, modeling, and human EEG data, but to date there have been no single-neuron or population recording studies that directly address this question. This is important because discerning between the two strategies could benefit from the spatial and temporal resolution afforded by single-unit electrophysiology, and because ultimately, we seek biological mechanism. Also, most tasks include a time delay between decision and confidence reporting (including the waiting-time assay; Lak et al., 2014; Stolyarova et al., 2019), potentially biasing participants toward a post-decisional strategy.

We circumvented these issues by training monkeys to report choice and confidence simultaneously in a response-time (RT) paradigm (‘peri-decision wagering’, peri-dw), building on previous work in human participants (Kiani et al., 2014a; van den Berg et al., 2016a). The goal was not to challenge post-decisional accounts of confidence, as this phase clearly plays a role in many tasks and natural situations. Rather, our task was designed to unveil whether a serial process is essential for a monkey's deliberation about both aspects of the decision, or if this process is fundamentally parallel. The approach facilitated evaluation of competing hypotheses on equal footing within the framework of bounded evidence accumulation, and to shed light on the underlying neural mechanism. Notably, the task measures choice, RT, and confidence on every trial, unlike ‘opt-out’ or uncertain-response paradigms (Kiani et al., 2009; Smith et al., 2012; Komura et al., 2013; Fetsch et al., 2014a; Li et al., 2023), allowing us to relate neural activity to all three measures on a trial-by-trial basis.

While monkeys performed the task, we recorded the activity of ensembles of neurons in the ventral portion of the lateral intraparietal area, LIPv (Shadlen & Kiani, 2013). Previous work has shown that LIPv (hereafter LIP) contains a representation of a decision variable (DV) that predicts choice and RT (Roitman & Shadlen, 2002), as well as confidence in an opt-out task (Kiani & Shadlen, 2009). Recent work using high-density probes found that the DV representation is observable on single trials (Steinemann et al., 2024) and thus is not an artifact of averaging (Latimer et al., 2015). Lastly, although one study failed to observe behavioral effects of LIP inactivation (Katz et al., 2016), at least four others support a causal role for this area in visuo-motor decisions (Hanks et al., 2006; Zhou & Freedman, 2019; Chen et al., 2020; Jeurissen et al., 2022). This makes LIP a sensible target for pursuing a population-level understanding of the dynamics of choice and confidence formation.

Using a combination of computational modeling and model-free analyses, we first show that a parallel model can explain all three features of the decision in this task, as suggested by previous work (Fetsch et al. 2014a; Kiani et al., 2014a). Then, using single-neuron analyses and population decoding, we show that the DV representation in LIP supports concurrent deliberation toward a

choice and associated wager. This work expands the toolkit for studying confidence in animal models and supports a role for LIP in decisions guided by an online estimate of confidence.

RESULTS

We recorded 407 neurons in area LIP of two rhesus monkeys (*Macaca mulatta*; 207 in monkey H, 200 in monkey G) while they performed the peri-dw task (Figure 1A). The saccade targets correspond to a motion direction decision (left or right) and a wager (high or low) on the correctness of that decision. Although behaviorally the task amounts to a choice among four options, we refer to the left-right decision as ‘choice’ and the high-low decision as ‘wager’ or ‘bet’, for simplicity and because the results support this interpretation. Monkeys were rewarded or penalized based on the conjunction of accuracy and wager (Figure 1B): a larger drop of juice for high vs. low bets when correct, and a time penalty for high bets when incorrect (no penalty for a low-bet error). As in previous studies (e.g., Roitman & Shadlen, 2002), monkeys showed greater accuracy (Figure 1B, left) and faster RTs (Figure 1B, middle) when the motion was strong compared to weak (coh near 0%). Motion strength also influenced wagering behavior in the expected manner, namely the probability of betting high increased with greater motion strength in either direction (Figure 1B, right). Importantly, the behavior shows that the low-bet option did not correspond to opting out of the motion decision: accuracy remained high on low-bet trials, and choice and RT still varied systematically with motion strength in a manner consistent with a deliberative process (see Model Fitting section below).

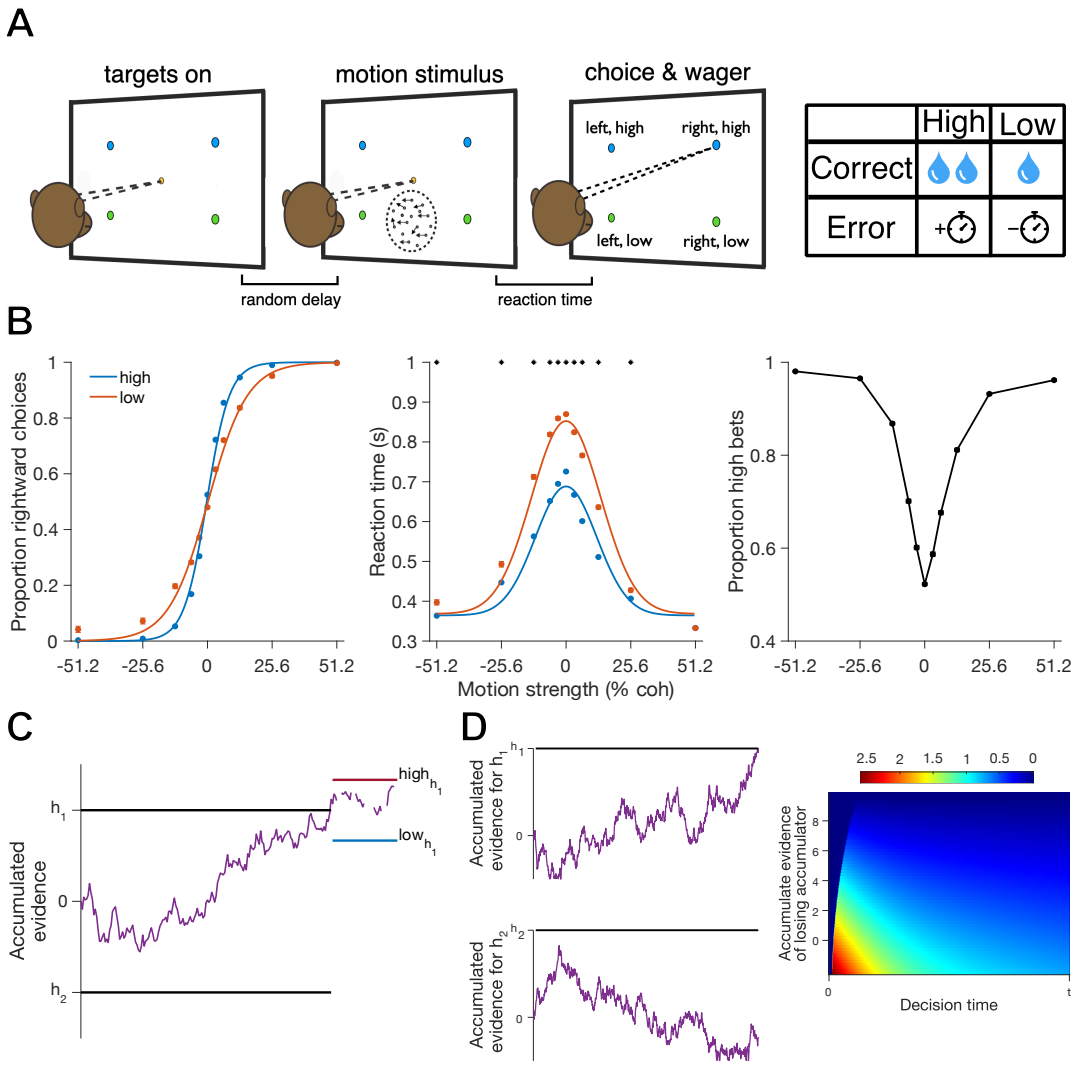

Figure 1. Task, behavioral performance, and model schematics. (A) After the monkey acquired fixation, four targets were presented, followed by a random-dot motion stimulus. At any time after stimulus onset the monkey could make a saccade to one of the targets to signal its choice and wager. The table at the right illustrates the possible outcomes for each trial: if correct, a high bet yielded a larger juice reward compared to a low bet, but if incorrect, a high bet incurred a 2-3 s time penalty assessed on the following trial's pre-stimulus fixation period. Low-bet errors were not penalized. (B) Performance on the peri-dw task pooled across two monkeys (N = 216 sessions, 202,689 trials, including sessions without neural recording). Each behavioral variable is plotted as a function of signed motion strength (%coh, negative=leftward, positive=rightward). Choice (proportion rightward) and RT functions are shown conditioned on wager (low=red, high=blue). Error bars (SE) are smaller than the data points. Smooth curves show logistic regression (choice) and Gaussian (RT) descriptive fits. (C) The serial model begins with a single accumulator with symmetric bounds (standard 1D drift-diffusion model). Arriving at one of the bounds terminates the primary decision (h_1 vs. h_2 = left vs. right in our task) and initiates a secondary process that accumulates evidence toward a 'high' or 'low' bound governing the wager. Crossing either of these bounds terminates the entire decision process and initiates the corresponding response. (D, left) The parallel model comprises two concurrent accumulators that are partially anti-correlated, allowing for variability across trials in the amount of evidence favoring the unchosen option. The first bound to be crossed (the 'winner') dictates the choice and decision time, whereas the losing accumulator dictates confidence by way of a mapping (right) between accumulated evidence and the log odds that the choice made was the correct one (color scale). Importantly, the mapping takes into consideration not only the amount of evidence but also the elapsed time.

Crucially for a behavioral assay of confidence, the monkeys' sensitivity was greater when betting high versus low (Figure 1B left, red vs. blue; $p<10^{-258}$, logistic regression). This was true even when controlling for variability in motion energy within each coherence level, by leveraging multiple repeats of the same random seed (Figure S1A; monkey H: $p<10^{-4}$, monkey G: $p<10^{-6}$; see Methods). Additionally, we fit a session-by-session logistic model separately for high and low bet trials, providing two distributions of weights relating motion strength to choice, and found that they were significantly different (Figure S2A; monkey H: $p<10^{-19}$, monkey G: $p<10^{-37}$). Both monkeys also showed faster RTs when betting high versus low, for all but the largest motion strengths (Figure 1B middle, red vs. blue; asterisks indicate $p<0.0045$ by t-test, Bonferroni corrected). Similarly to choices, we fit a Gaussian function to the average RT as a function of motion strength for each session, separately for high and low bet trials. The amplitude parameters of the fitted Gaussians were greater for low vs. high bets (Figure S2B; monkey H: $p<10^{-7}$, monkey G: $p<10^{-25}$, paired t-test), indicating greater modulation of RT by motion strength when the monkey indicated low confidence. Lastly, we examined wagering behavior as a function of RT (Figure 1B right), separately for each individual motion strength (Figure S3A,B). For most motion strengths, the monkeys bet high less often for longer RTs (Figure S3B; $p<.0085$ for both monkeys for every coherence except 51.2%, Cochran-Armitage trend test with Bonferroni correction). This pattern is remarkably similar to that observed in humans on a similar task (Kiani et al., 2014a), where the results argued that confidence depends on both evidence strength and elapsed time. As in that previous study, the pattern remained significant when controlling for variability in motion energy across trials of a given coherence (Figure S3D; monkey H: $p=0.0003$, monkey G: $p=0.0056$, interaction term between motion energy and RT quintile using ANCOVA).

An inverse relationship between response time and confidence is a classic psychophysical result (Henmon, 1911; Kellogg, 1931; Audley, 1960), replicated in more recent human work (e.g. Kiani et al., 2014a; Desender et al., 2021; Dou et al., 2024). Observing it in monkeys, for the first time to our knowledge, supports the notion that the peri-dw assay is a valid measure of confidence, and it is consistent with a family of accumulator models as addressed below. Note that we

sometimes refer to the wager (or proportion of trials with a high bet) as ‘confidence’ for simplicity, acknowledging that confidence is a latent cognitive variable and any behavioral measure is bound to be imperfect.

Model-free analyses of behavior suggest temporally overlapping choice and confidence computations

Although the choice and wager were indicated with a single eye movement, this does not necessitate simultaneity in the processing of evidence. Different temporal windows of the stimulus could covertly be used to support the two elements of the decision, which would then only be reported when both were resolved. To test whether monkeys use a consistent serial strategy (resolving choice first and then confidence, or vice versa) we calculated the influence of stimulus fluctuations on choice and confidence as a function of time (psychophysical kernels; Kiani et al., 2008; Nienborg & Cumming, 2009; Zylberberg et al., 2012). Briefly, we quantified the motion energy for each trial and video frame by convolving the random-dot pattern with two pairs of spatiotemporal filters aligned to leftward and rightward motion (Adelson & Bergen, 1985). We then partitioned trials by outcome (choice and wager) and plotted the average relative motion energy (residuals) for each outcome as a function of time.

Psychophysical kernels for choice are plotted in Figure 2A. Rightward choices were preceded by more rightward motion energy throughout most of the trial (red line), and the same was true for leftward choices and leftward motion (blue). The kernels for right and left choice began to separate about 100 ms after motion onset and remained so until ~100 ms before saccade initiation. This clear separation, aligned on both motion onset and saccade, suggests that the monkeys used essentially the entire stimulus epoch to decide motion direction. For confidence, we calculated the kernels by taking the difference between the motion energy time series for high and low bets associated with a specific choice (van den Berg, 2016a). For high minus low wager on rightward choice trials, motion energy

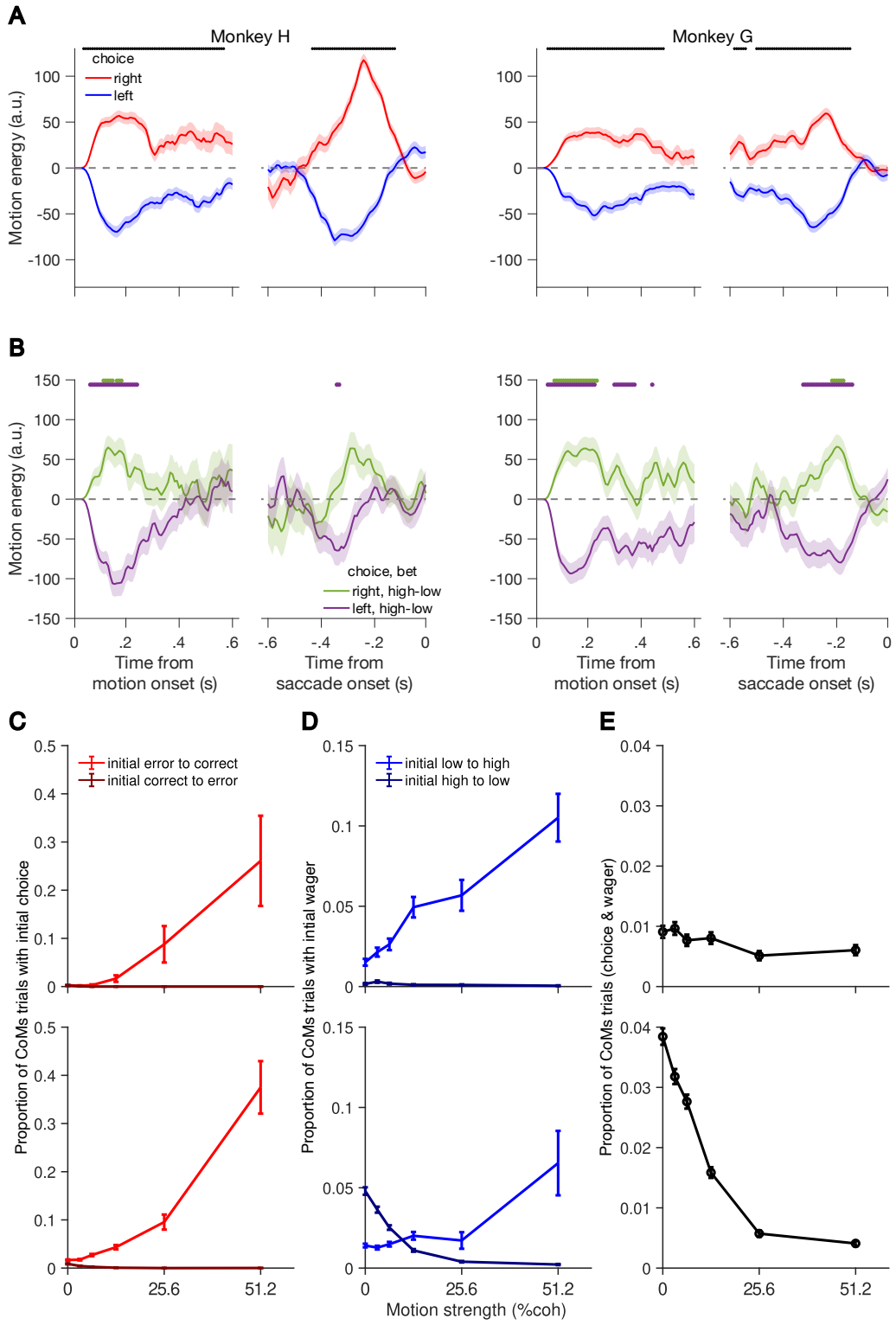

Figure 2. Psychophysical kernels and changes-of-mind for choice and wager. (A) Motion energy profiles conditioned on right and left choices (red and blue, respectively), aligned to motion onset and saccade onset. Shaded regions indicate SEM. Left column is for monkey H and the right for monkey G. Black line at top indicates when right and left traces were significantly different from each other ($p<0.05$, t-test with Šidák Correction for 140 frames). (B) Confidence kernels computed as the difference in motion energy between right-high and right-low choices (green), and the difference between left-high and left-low choices (purple), aligned to the same events as (A). Colored lines at top of graph indicate when the corresponding traces were significantly different from zero ($p<0.05$, t-test with Šidák correction).

157

values were above zero, indicating an excess of rightward motion on high-bet choices compared to low-bet choices (Figure 2B, green). Similarly, for left choice trials the difference in motion energy was below zero, indicating more leftward motion on high vs. low bets (Figure 2B, purple). This analysis shows that both early and late motion evidence is leveraged to inform confidence, for both monkeys. Comparison of the traces in Figure 2A vs. 2B might suggest that the utilization of the stimulus for confidence does not identically overlap with choice, especially for monkey H. However, the substantial overlap does appear to rule out a consistent temporal segregation, such as an obligatory post-decision mechanism for confidence. Of course, psychophysical kernels rely on trial averages and cannot resolve dynamics of individual decisions; we partly address this shortcoming below using neural recordings.

When decisions are reported with an arm movement (Resulaj et al., 2009; van den Berg et al., 2016a), human subjects occasionally alter their reach trajectory in a manner that suggests a 'change of mind' (CoM) based on continued processing of evidence after movement initiation. Saccadic choices are generally considered incompatible with CoMs because of their speed and ballistic nature (Resulaj et al., 2009; but see McPeek et al., 2000; Caspi et al., 2004), but we were nevertheless able to identify a small subset of trials with multiple saccades in quick succession that revealed putative CoMs (see Methods). These trials showed characteristic features of CoMs (Figure 2C-E), including greater frequency on difficult vs. easy trials (Figure 2E; monkey H: $p = 0.0076$; monkey G: $p<10^{-6}$, Cochran-Armitage test). Changes from incorrect to correct were more likely when motion strength was high (Figure 2C red line; monkey H: $p < 0.0001$; monkey G: $p < 10^{-5}$), whereas correct-to-error CoMs, occurring sparingly, were more likely when motion strength was low (Figure 2C, dark red line; monkey H: $p < 0.0012$ monkey G: $p < 10^{-5}$). We also observed changes from low to high confidence, which for one monkey were more frequent with greater motion strength (Figure 2D blue line; monkey H: $p < 10^{-60}$ monkey G: $p = 0.944$), as shown previously in humans (van den Berg et al., 2016a). The presence of CoMs and changes of confidence (sometimes both on the same trial) suggests that both dimensions of the decision were subject to revision at the time of the initial saccade. This argues against a strictly serial process,

although it also highlights a potential window of post-decisional processing, even for saccadic decisions (McPeek et al., 2000; Caspi et al., 2004).

Model fitting favors parallel deliberation for choice and confidence

Previous studies have typically focused on either a serial (Pleskac & Busemeyer, 2010; Moran et al., 2015) or a parallel (Vickers, 1979; Shadlen & Kiani, 2013) framework for confidence. Thus, explicit comparisons of these models have generally been qualitative because of variations in task design across studies (but see Shekhar & Rahnev, 2024 for a comprehensive evaluation). Here we provide a quantitative comparison of the two classes of model applied to the same data from the peri-dw task. Our exemplar of the serial strategy is a one-dimensional drift-diffusion model (DDM) with post-decision accumulation for confidence (Moran et al., 2015; Herregods et al., 2023; Figure 1C). The choice is determined by which of two primary decision boundaries is crossed, and the wager by a second set of bounds symmetric around the primary (winning) bound. The observed RT is the sum of the time taken to reach both bounds, plus non-decision time. Our implementation of a parallel strategy is a two-dimensional DDM, where a common accumulation epoch governs both choice and confidence (Kiani et al., 2014a; van den Berg et al., 2016; Figure 2B). Here there are two accumulators integrating evidence in support of the two alternatives, equivalent to an anticorrelated race. Two accumulators are necessary because the winning accumulator determines the choice and RT while the losing accumulator is leveraged to compute confidence. In contrast, the serial model uses two separate epochs of accumulation to inform choice and confidence, and therefore requires only a single accumulator. Note that the key distinction is serial vs. parallel, not 1D vs. 2D; we simply chose these as exemplars based on the leading candidates in the literature.

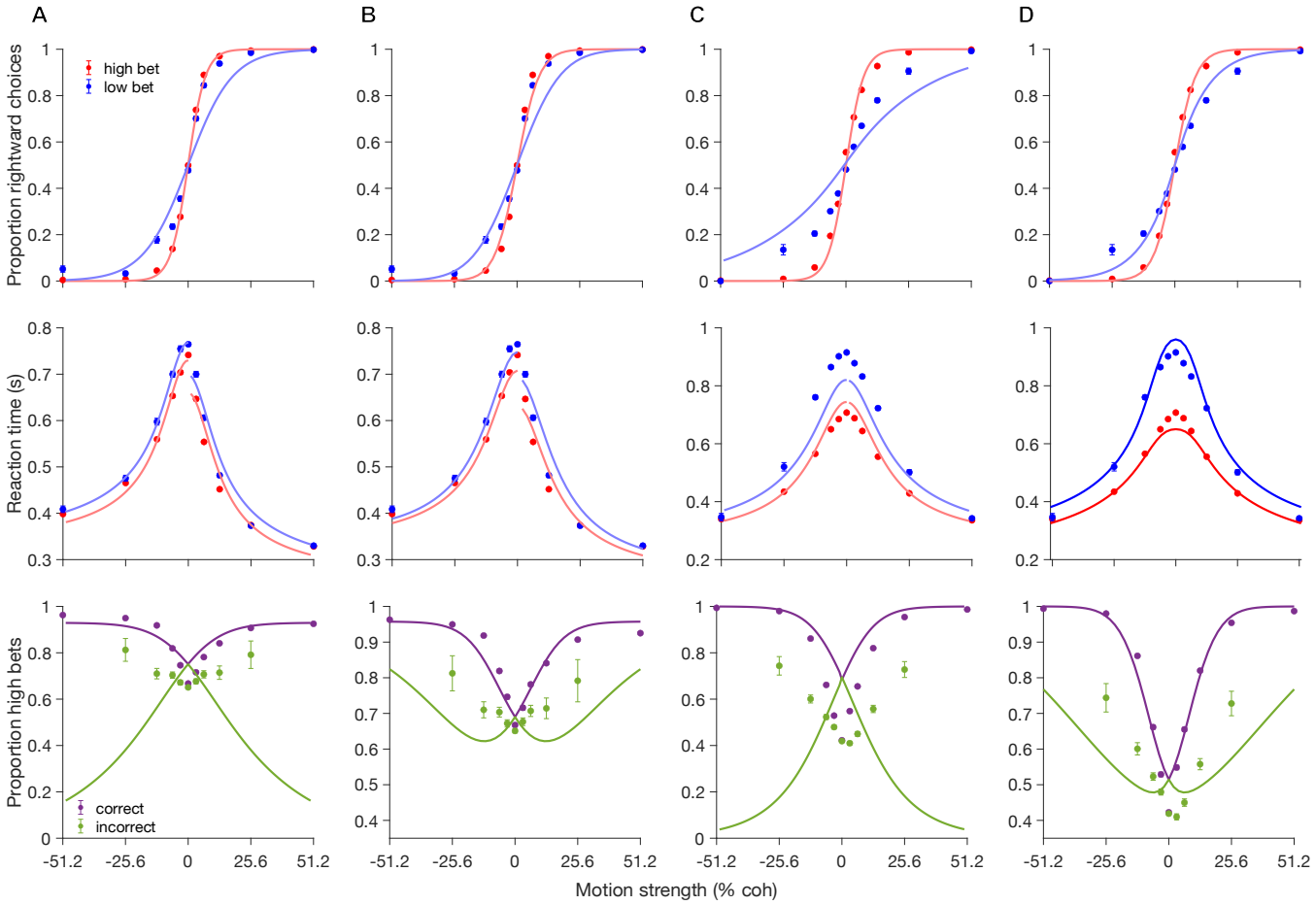

Figure 3. Comparison of serial and parallel model fits. The left two columns show results from monkey H and the right two columns monkey G. Within these two columns the left column shows fits to the serial model and the right column the parallel model. (A) Proportion of rightward choices as a function of motion strength (% coh), conditioned on high and low bet trials (red and blue, respectively). The solid dots are empirical data points and solid lines are model fits. (B) Similar to (A) but for mean reaction time. (C) Proportion of high bets as a function of motion strength (% coh), conditioned on correct and error trials (magenta and green, respectively). The fitting procedure used for these panels differs from the previous ones. They are fitted using only correct trials, and the green traces illustrate the resulting prediction for error trials.

The curves in Figure 3 and Figure S4 are fits to the serial (A and C) and parallel models (B and D) for the two monkeys. Both models capture the general trends in choice, RT, and wager across motion strengths, at least when pooled across correct/incorrect and high/low wager trials (Figure S4), and both qualitatively predict greater sensitivity and faster RTs for high vs. low wager. However, in monkey G the serial model was unable to capture the large difference in reaction time between high and low bet trials (Figure 3C, middle). The more subtle RT difference in monkey H was well handled by the serial model (Figure 3A, middle), but this came at the cost of a slightly poorer fit to the wager-conditioned choice data (Figure 3A vs. B, top). Thus, the behavioral

differences between the two animals served to highlight complementary strengths of the parallel model: its ability to explain a broader range of confidence-RT relationships while also tolerating a more subtle increase in sensitivity on high vs. low-bet trials. Quantitative Model comparison supported these observations: Bayesian information criterion (BIC) was lower for the parallel vs. the serial model (monkey H: 8.24×10^5 (parallel) < 8.79×10^5 (serial), monkey G: 1.21×10^6 (parallel) < 1.30×10^6 (serial)).

More strikingly, the serial model failed qualitatively to reproduce the confidence pattern on error trials (Figure 3, bottom row). In general, confidence increases as a function of stimulus strength for correct choices—which makes intuitive sense—but empirically it often *decreases* with stimulus strength on incorrect trials. This characteristic 'X-shape' pattern has been proposed as a statistical signature of confidence in behavior (Sanders et al., 2016) and brain activity (Kepecs et al., 2008; Rolls et al., 2010; Komura et al., 2013; Bang et al., 2020). In contrast, we (Figure 3, bottom row) and others (Kiani et al., 2014a; van den Berg et al., 2016a) observed that confidence increases with motion strength even for errors. As noted previously (Kiani et al., 2014a; Fetsch et al., 2014b; Desender et al., 2021; Khalvati et al., 2021), these conflicting findings can be explained by the relative timing of choice vs. confidence. Resolving choice first followed by confidence later allows for revision of the confidence judgment upon further deliberation. Incorrect choices when the stimulus was strong are more likely to undergo such revision, as further processing reveals that the initial choice was incorrect. However, when choice and confidence are reported simultaneously there is little or no time for revision, and thus confidence on error trials either increases or remains flat as a function of stimulus strength (Kiani et al., 2014a; van den Berg, 2016a, Desender et al., 2021). For this reason, the serial model tested here simply cannot reproduce the error-trial confidence pattern we observed empirically (note that the green curves in Figure 3, bottom, are a prediction, not a fit). In summary, although basic behavioral patterns are reasonably well predicted by the serial model, model comparison favors a process where evidence is accumulated in parallel for constructing a decision and associated level of confidence.

LIP neurons show signatures of concurrent accumulation

Putative DV representations have been observed in several subcortical and cortical areas, including LIP (Shadlen & Kiani, 2013), as well as in aggregate signals observable with noninvasive methods in humans (O'Connell & Kelly, 2021). Although the widespread 'ramping' pattern of activity does not uniquely identify a process of evidence accumulation, a large body of work supports the assertion that LIP activity reflects such a process during the random-dot motion task. We reasoned that, if choice and confidence were resolved concurrently during motion viewing (parallel model), the ramping activity should begin to predict both dimensions of the eventual saccade at the same time, classically around 200 ms after motion onset (Roitman & Shadlen, 2002). Alternatively, if choice were to be deliberated first followed by confidence later (serial model), this temporal separation should be evident in the divergence point of neural activity traces conditioned on the four outcomes.

These traces are shown in Figure 4A for four example neurons. The highest firing rate corresponds to choices made into the response field (RF) of the neuron, which was almost always in the left (contralateral) hemifield but was equally likely to overlap the high or low wager target. The relative ordering of the remaining three traces differs across neurons, possibly due to idiosyncratic RF properties or nonspatial choice or confidence signals. The key observation is that the activity preceding saccades to the preferred wager target (low or high) diverges from the activity for the other wager target (high or low) at about the same time as it diverges from the traces for ipsilateral choice (right-low and right-high). This pattern is present in each example neuron as well as in the population averages (Figure 4B,C). There is no evidence that ramping activity consistently predicts the left-right choice prior to the high-low one (or vice versa), as expected under a serial model. Instead, to the extent the activity reflects accumulation of evidence favoring the target in the RF (see below and Discussion), the results support a model in which such accumulation underlies concurrent deliberation toward a choice and confidence judgment.

To dig deeper into the nature of the observed ramping signals, we tested for statistical signatures of a noisy accumulation process (Churchland et al., 2011; de Lafuente et al., 2015;

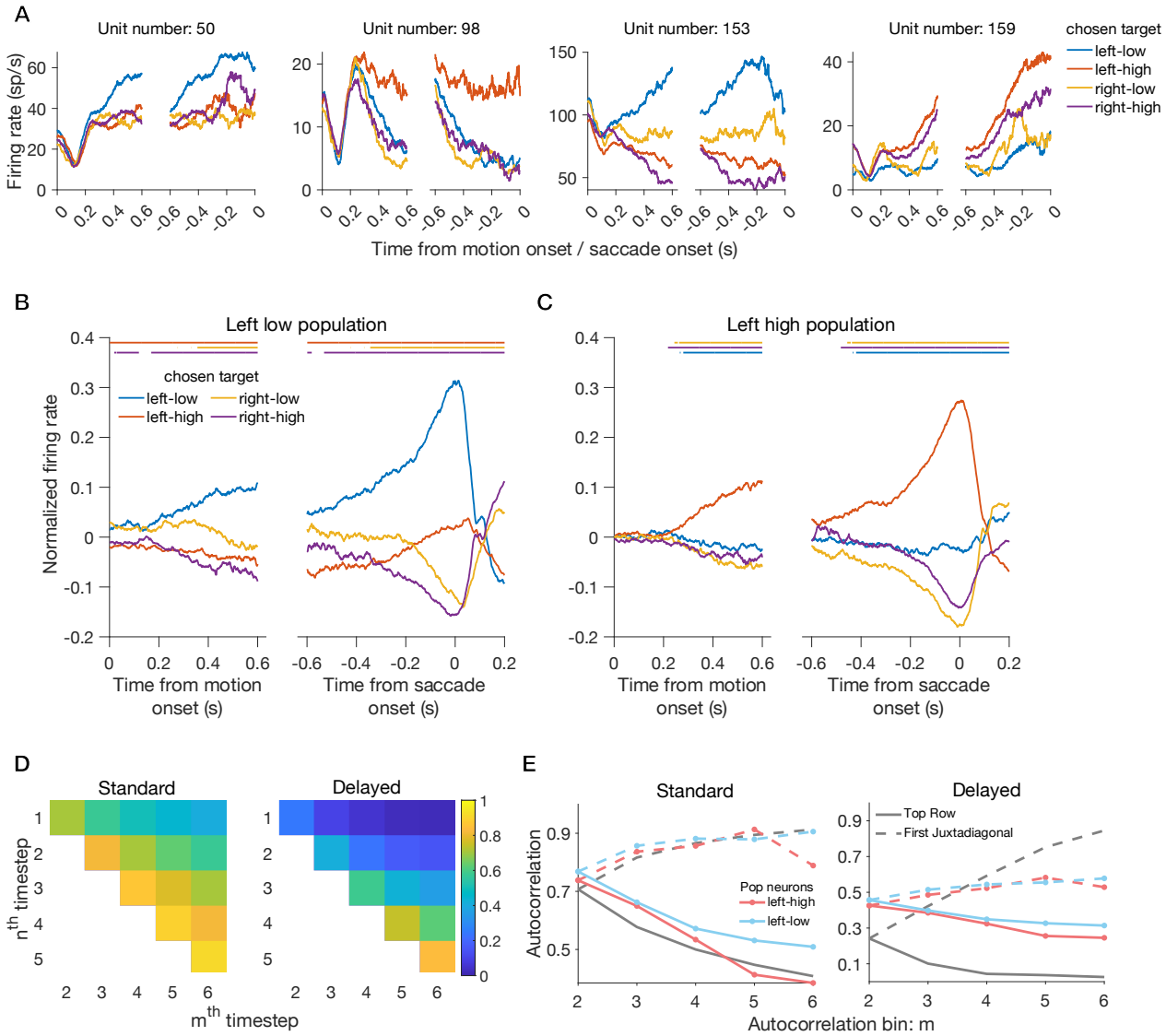

Figure 4. Temporal properties of neural activity in LIP during the peri-dw task. (A) Firing rate of example units split by choice/wager outcome, aligned to motion onset and saccade onset. (B) Population average firing rate (normalized) for neurons with an RF overlapping the left-low target. Colored bars at top indicate when the corresponding trace is significantly below the trace for choices into the RF. (C) Same as B but for left-high neurons. Only low-coherence trials (0, +/- 3.2%, +/- 6.4%) are included in A-C. (D) Theoretical autocorrelation matrix of a standard ideal accumulation process (left) and a delayed accumulator (right; see Methods). (E) Projection of theoretical autocorrelations for top row (gray dotted) and first juxtadiagonal (gray solid) along with the corresponding data after fitting the phi parameter (Methods). Red and blue represent the two populations shown in B,C, pooled data from both monkeys.

Shushruth et al., 2018; Steinemann et al., 2024): increasing variance of the latent rate as a function of time (variance of conditional expectation, VarCE) and a characteristic autocorrelation pattern in this latent signal (correlation of conditional expectation, CorCE). These signatures follow from considering neuronal spiking as a doubly stochastic process described by a Poisson distribution with an underlying rate that varies within and across trials. For stochastic bounded

accumulation, VarCE should increase linearly with time during deliberation, then decrease near the saccade as the bound is reached. Second, correlations between firing rates at adjacent time points (t and t+1) should increase for pairs of bins further out in time from stimulus onset (t+n and t+n+1, as n increases). Third, the correlation between two time points should decrease as the distance between them increases (t and t+1 versus t and t+5; Figure 4D, left & Figure 4E, left).

The results supported all three predictions. After 200 ms following motion onset, VarCE shows a roughly linear increase for at least the next 400 ms (Figure S5, left; left-high neurons, monkey H: $p < 10^{-7}$, monkey G: $p < 10^{-13}$; left-low neurons, monkey H: $p < 10^{-19}$, monkey G: $p = 0.0014$; linear regression), then decreases near saccade initiation (Figure S5, right; left-high neurons, monkey H: $p = 0.015$, monkey G: $p = 0.0016$; left-low neurons, monkey H: $p < 10^{-4}$, monkey G: $p < 10^{-7}$). For CorCE, the results from both monkeys were reasonably well matched to the predictions (Figure 4E, left; monkey G: $R^2 = 0.74$ and 0.79 for left-high and left-low neurons, respectively; monkey H: $R^2 = 0.69$ and 0.79, respectively).

These dynamics in variance and autocorrelation are consistent with an underlying neuronal process that reflects accumulation, and are not easily explained by alternative accounts of LIP ramping activity such as a gradual shift of attention or simple movement preparation. Critically, the patterns were present over the same time window in both the high- and low-preferring populations. This argues against a form of the serial model in which choice is initially resolved by considering only one pair of targets, followed by a shift to the other pair after some time has elapsed. We explicitly tested this by computing the expected autocorrelation under a simulated process where integration is delayed by a random amount of time. Fitting such a model to the data generated a qualitative and significant mismatch with the CorCE prediction (Figure 4D, right & Figure 4E, right; monkey H: $p < 0.008$ for both populations, monkey G: $p < 0.012$ for both populations). Taken together, the results favor a parallel model in which deliberation occurs simultaneously between both the high and low pairs of targets. What remains to be tested is whether and when these accumulation signals are predictive of the monkey's choice and wager on individual trials.

Single-trial decoding supports a link between choice and confidence signals

Thus far most of our analyses have relied on trial averages, potentially obscuring the dynamics of individual decisions. We therefore turned to a population decoding approach (Kiani et al., 2014b; Kaufman et al., 2015; Peixoto et al., 2021; Steinemann et al., 2024), leveraging the simultaneous recording of neural ensembles to more directly address the question of parallel vs. serial deliberation. We trained two logistic classifiers, one for the binary choice and another for the binary wager, using the spike counts from the neurons recorded during each session (mean = 14 units/session). The analysis determines the linear weight for each neuron that maximizes the probability of predicting the observed choice or wager (Methods). We can then extract a 'model decision variable' which is simply the log odds of a particular choice or wager, according to the decoder, based on the population spike counts up to time *t* on a given trial (Kiani et al., 2014b; Peixoto et al., 2021). This quantity is also known as prediction strength or certainty. To facilitate comparison of the two decoders, we collapsed across choices such that positive values of the model DV correspond to correct prediction of the choice, rather than exhibiting symmetry around zero as in previous work.

For both monkeys, the model DV ramped up starting about 200 ms after motion onset (Figure 5A), as shown previously. The DV dynamics differed for the two animals, but both exhibited a ramping slope that depended on motion strength (monkey H: $p < 0.001$, monkey G: $p < 10^{-4}$, linear regression). Cross-validated prediction accuracy also ramped up starting at this time, and critically, did so simultaneously for both the choice and confidence decoders (Figure 5B). Horizontal bars at the top of the plot indicate significant differences from chance level, which occurred for both decoders starting around 200 ms and persisted until saccade onset. At their peaks, both performed well above chance on the test set (Figure 5B), but a notable difference is the timing of the peaks, which for choice is just before saccade onset and for wager is slightly after the saccade (Figure 5B, right). The black bar in Figure 5B indicates a period near the saccade with

327 a significant difference in prediction accuracy between the decoders, highlighting an epoch where

328 choice accuracy peaks but wager accuracy continues to increase.

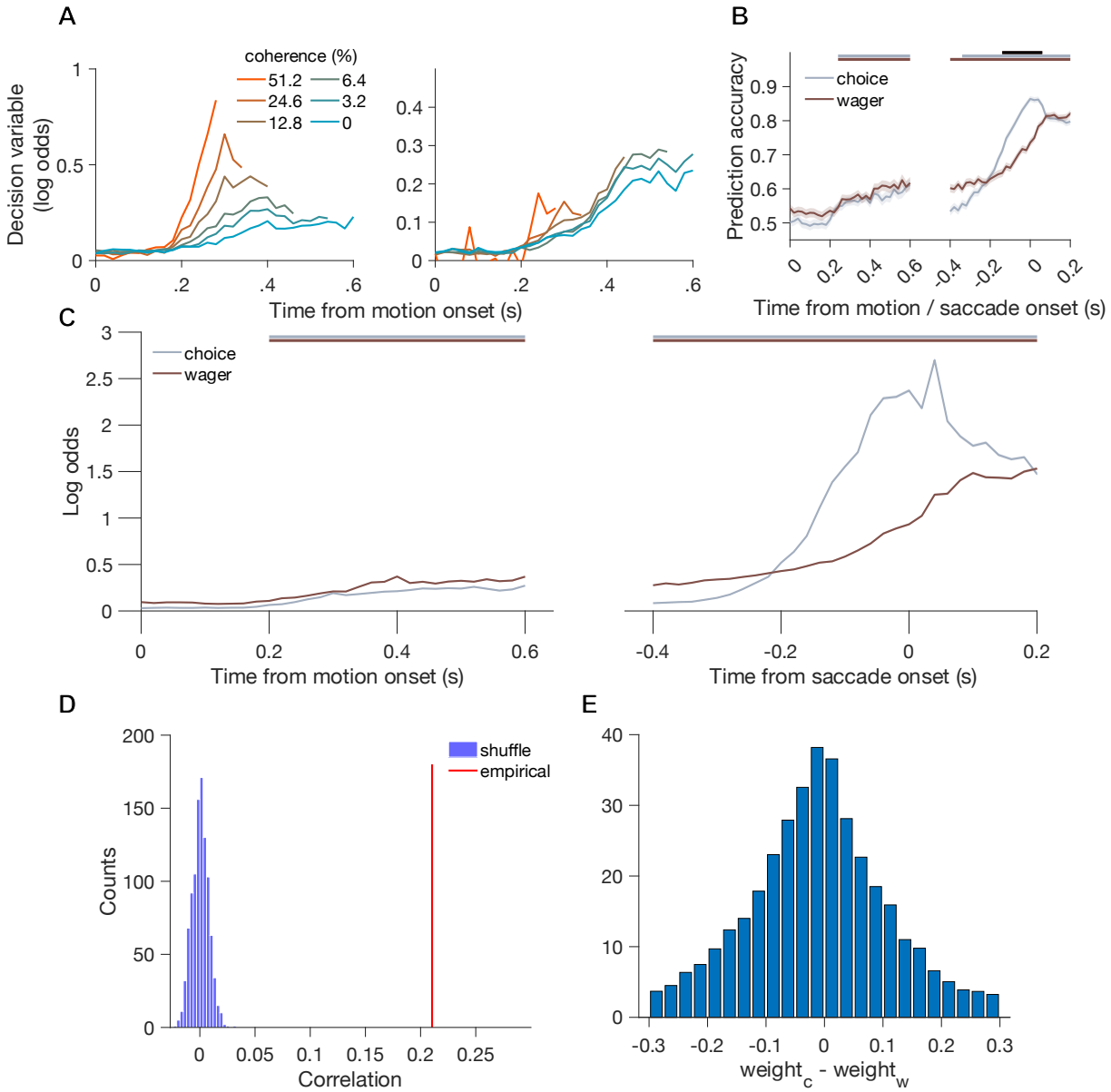

Figure 5. Population decoders support concurrent readout of choice and wager from a unimodal population. (A) Log odds ('model decision variable') quantifying prediction strength ('certainty') for the choice decoder as a function of time, aligned on motion onset and conditioned on motion strength. Left graph is monkey H and the right graph is monkey G. (B) Prediction accuracy (proportion correct binary classification in the test set), for both the choice (gray) and wager (brown) decoders, as a function of time and aligned to motion onset and saccade onset. Shaded regions around the traces indicate SEM. Gray and brown bars at top indicate when accuracy for the corresponding decoder was significantly greater than chance. Black bar indicates when prediction accuracy was significantly different for choice vs. wager. (C) Log odds for both choice and wager decoders as a function of time and aligned to motion onset and saccade onset. Line color, error shading, and significance bars are similar to 5B. (D) Correlation between the magnitude of choice and wager weights (red line) compared to the value expected by chance (shuffled data). (E) Histogram of the difference between choice and wager decoder weights.

329 We next confirmed that this temporal offset is not simply a product of the binary
330 classification method but also reflects graded prediction strength (log odds or model DV). The log
331 odds for both choice and wager were significantly different from chance after 200 ms and persisted
332 until saccade onset (Figure 5C). As with prediction accuracy, the initial peak for choice is near
333 saccade onset while the peak for the wager occurs about 100 ms after the saccade. We speculate
334 that this late peak may be related to continued evaluation of evidence for a possible CoM (see
335 below), although we cannot exclude the possibility of a signal related to expectation of outcome
336 (reward or penalty).
337 The results in Figure 5B and 5C offer clear support for parallel deliberation of choice and
338 wager, refuting a temporal bottleneck associated with a serial strategy. However, a different form of
339 bottleneck might exist within the population of neurons that manifests as distinct subsets of
340 neurons supporting choice and wager. The fitted decoder weights used to predict choice and
341 wager serve as tools for addressing this question. To summarize the relationship between the
342 choice and wager weights we converted them to an absolute magnitude and performed a
343 Pearson's correlation (Figure 5D) and a difference test (Figure 5E). For both analyses the weight
344 time series were marginalized over the accumulation period (Methods). The correlation across the
345 population was modest (monkey H: $r = 0.18$ & monkey G: $r = 0.21$) but highly significant by
346 permutation test ($p << 0.05$ for both monkeys). Additionally, the distribution of weight differences
347 was unimodal (Figure 5E, Hartigan's dip test, $p > 0.9$ for both monkeys individually), suggesting a
348 continuum of contributions to choice and confidence across the population and not two distinct
349 subpopulations contributing to one or the other.
350 As mentioned, one way to describe the model DV is as a graded level of certainty (Kiani et
351 al., 2014b) or even "confidence" (Peixoto et al., 2019) in the prediction of behavior by neural
352 activity. Although these are just labels (see also Pouget et al., 2016), we wondered whether the
353 level of certainty of the choice decoder might predict the binary classification by the wager decoder
354 on a trial-by-trial basis. To test this, we partitioned trials according to whether they were classified
355 as high or low bets by the wager decoder, by calculating the mean P(High) for the wager decoder

356 in the period immediately surrounding the saccade. Values above (below) 0.5 indicated a decoded-
357 high (decoded-low) wager. We then averaged the model DV from the choice decoder, using only
358 0% coherence trials, and found that it was higher for decoded-high vs. decoded-low trials (Figure
359 6A). This indicates that the strength with which LIP predicts the upcoming choice covaries with the
360 degree to which the same population predicts a high bet, consistent with a tight link between
361 choice and confidence signals. This is surprising given that the spatial/motor dimensions for choice
362 and wager are orthogonal.

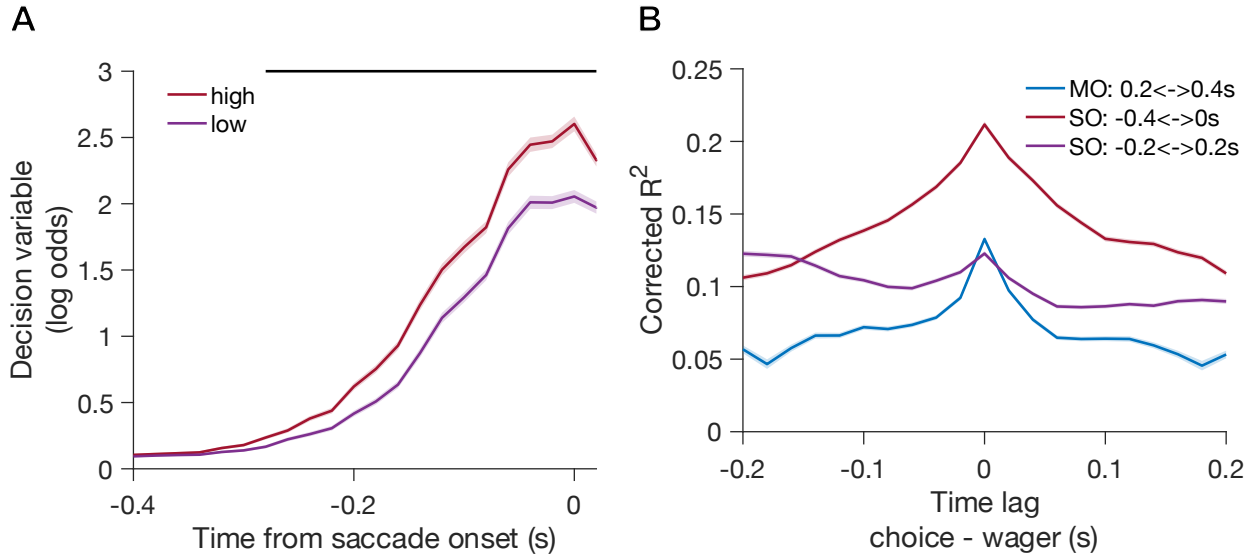

Figure 6. Predictive relationship between choice and wager decoders. (A) Log odds of choice decoder as a function of time, aligned to saccade onset. Red and purple lines are trials separated by whether the wager decoder predicted a high or low bet. Shaded regions are SEM. Black bar at top indicates a significant difference between the traces. (B) Corrected R^2 values from a linear regression relating trial-by-trial transformed choice decoding strength and wager decoder probability (P(High)), as a function of time lag. The blue, red, and purple traces comprise different time windows either aligned to motion onset (MO) or saccade onset (SO). Standard error bars are included as shaded regions near the lines (barely visible). Negative values on the x-axis indicate wager decoder values were regressed with choice decoder values that were x seconds in the past, and vice versa for positive values.

363 With the established link between decoded choice certainty and wager prediction, we can
364 now reexamine the temporal differences shown in Figure 5B & 5C, specifically by asking when
365 choice decoding strength best predicts the wager decoder's probability of a high bet on a given
366 trial. We fit a linear regression relating choice decoder strength (Methods) at time t to decoded
367 wager probability at time $t + \Delta t$, where Δt ranges from +/- 200 ms. We found that, during the
368 deliberation phase (200-600 ms after motion onset and 400-0 ms before saccade onset) the best
369 prediction between choice strength and wager probability was at zero time lag (Figure 6B, R^2 = .13

& .21 blue and red lines, respectively). Interestingly, the period centered around the saccade gave rise to two peaks, one at zero lag (R^2 = 0.12) and one at a lag of -0.2 s (wager lagging behind choice; R^2 = 0.125). This suggests that the updating of the confidence signal around the decision time instead reflects a procrastinated element (perhaps evidence, or a bias signal) that was integrated into the decision variable earlier in time.

DISCUSSION

The neural mechanisms underlying metacognition have become tantalizingly more accessible over the past two decades through the development of behavioral assays of confidence in nonhuman animals (Hampton, 2001; Kiani et al., 2009; Middlebrooks & Sommer, 2011; Smith et al., 2012; Kepecs & Mainen, 2012). A longstanding goal is to connect the rich psychological literature on process models for confidence with their implementation at the level of neural populations and circuits. One approach considers how decision accuracy, speed, and confidence can be jointly explained within the framework of bounded evidence accumulation (Kiani & Shadlen, 2009; Pleskac & Busemeyer, 2010; Fetsch et al., 2014b), an idea presaged by Vickers' balance-of-evidence hypothesis (Vickers, 1979). Such a framework is motivated by the critical role of response time in psychophysical theory and experiment (Luce, 1986) and its strong empirical link to confidence going back at least a century.

Yet embracing a dynamic model still leaves open questions about the temporal evolution of choice and confidence in the brain. Several authors have emphasized post-decisional processing (Baranski & Petrusic, 1998), formalized by serial models in which evidence is integrated for confidence only after termination of the primary decision (Pleskac & Busemeyer, 2010; Moran et al., 2015; Herregods et al., 2023). This idea follows naturally from the definition of confidence as the expected probability correct conditioned on a choice (Pouget et al., 2016), and it is sensible to exploit additional information acquired (or generated internally) after commitment to the choice if the behavioral context allows it. However, in many other settings it would seem advantageous to compute a provisional degree of confidence while the decision is still being formed. Decisions are

commonly expressed as motor actions, executed with a degree of vigor (or caution) that depends on confidence or expected reward (Shadmehr et al., 2019). Also, deliberation itself is a significant cost (Drugowitsch et al., 2012), so it could be less efficient to extend the deliberation period for confidence. Lastly, an online prediction of accuracy facilitates rapid decision sequences (van den Berg et al., 2016b; Lisi et al., 2020; Zylberberg, 2021) and strategic modulations of the decision process (Balsdon et al., 2020). Recent human studies (Dotan et al., 2018; Balsdon et al., 2020; Li et al., 2023) offer intriguing evidence for parallel computation of confidence (or at least *certainty*; Pouget et al., 2016) during decision formation, further supported by electroencephalography (Gherman & Philiastides, 2015; Balsdon et al., 2021; Dou et al., 2024) and transcranial magnetic stimulation (Xue et al., 2023). How this might be implemented at the level of neuronal populations remains unclear and requires a suitable paradigm for nonhuman animals. Here we establish such a paradigm and provide evidence that monkeys (and LIP neurons) can accumulate samples of evidence concurrently to guide a single motor action corresponding to a choice and degree of confidence therein.

Behavioral evidence for concurrent deliberation

First we validated the peri-decision wager as a measure of confidence *per se*, which should covary with accuracy even when conditioned on the stimulus. If a behavioral report only predicts accuracy across levels of stimulus strength or difficulty, it could reflect an associative process of categorization (Smith et al., 2012), or more subtly, an estimate of difficulty itself (Löffler et al., 2023). We sought to allay these concerns by including a set of identical-seed trials, in which the random dot movie was identical up to the initiation of the saccade. Controlling for stimulus motion, the accuracy for high bets was still significantly higher than for low bets, for both monkeys (Figure S1A). This suggests that the wager reflects, in part, variability in the internal representation of evidence and/or the reliability of the decision process, including aspects that could fall under the rubric of attention.

Psychophysical kernels indicated substantial overlap in the stimulus period contributing to choice and wager, as shown previously in humans (Zylberberg et al., 2012; van den Berg et al., 2016a), ruling out a strategy where confidence is systematically assessed after (or before) choice formation. Additionally, a small subset of trials showed evidence for changes of mind (CoMs): perhaps surprisingly for a saccade task, the monkeys occasionally reversed their initial wager, choice, or both, implying that both were amenable to revision at the time of initial commitment. These two lines of evidence suggest that deliberation for choice and confidence occurs in overlapping time windows, even extending after the stimulus has been extinguished.

Corroboration of a parallel strategy was provided by the superior fits and predictions from our parallel model compared to the serial model (Figure 3). We then used the model fits to make predictions for the probability of betting high when a trial outcome is correct vs. incorrect. This aspect of the behavior strongly argued against the serial strategy, which invariably predicts an 'X-shape' pattern (decreasing confidence with greater evidence strength on error trials) that was absent in our data. We did not explicitly test a hybrid model where parallel accumulation is followed by a period of extra accumulation for confidence (Desender et al., 2021; Maniscalco et al., 2021), but this too predicts an X-shape, depending on the duration of the second epoch (Desender et al., 2021). Other behavioral evidence (Kiani et al., 2014a) and model simulations (Khalvati et al., 2021) similarly appeal to the temporal domain to explain the X-shape, contrary to the conventional explanation based on signal-detection theory (SDT; Treisman & Faulkner, 1984; Kepecs et al., 2008) or more general Bayesian formulations (Sanders et al., 2016; Adler & Ma, 2018). The distinction matters because neural correlates of the X-shape are considered diagnostic of confidence signals in the brain (Kepecs et al., 2008; Rolls et al., 2010; Komura et al., 2013; Bang et al., 2020), which could be misleading if the relative timing of choice and confidence computations are unknown. In general, however, the temporal explanation does not invalidate SDT as a useful framework, nor its extension to an estimate of decision reliability ('meta-uncertainty'; Boundy-Singer et al., 2022). In fact an exciting future direction would be to link the neural

signatures of meta-uncertainty computations in visual cortex (Boundy-Singer et al., 2024b) to the temporal dynamics of confidence formation in decision-related areas.

Neural correlates of concurrent deliberation

Individual LIP neurons showed heterogeneous dynamics (Figure 4A), but when combined into populations with RFs overlapping either the left high or left low target, the averages revealed accumulation-like responses with interpretable timing relationships for the four outcomes. In essence, the ramping activity began to predict the wager at about the same time that it began to predict the left-right decision. Using a theoretical and simulated autocorrelation structure for a standard and delayed drift diffusion process, we found that both populations show the statistical signatures of accumulation (VarCE and CorCE) from 200 ms after motion onset, inconsistent with one form of serial strategy in which the decision process first considers only the low or high pair of targets then covertly switches to the other pair. Note that previous studies examined these signatures on simple 2-choice paradigms, so it was unclear whether LIP neurons in our 4-target task would conform to the same theoretical autocorrelation structure. It will be useful to generate specific predictions for the variance and autocorrelation of neural signals in more complex paradigms, including multi-alternative (Churchland et al., 2008) and multi-attribute perceptual and economic decisions (Kang et al., 2021; Sampson et al., 2023).

To uncover latent decision and confidence signals on individual trials, we constructed logistic classifiers for choice and wager based on the neural population activity. As shown previously, the choice decoder displayed a graded prediction (the ‘model DV’) that was linearly dependent on coherence, reminiscent of classic LIP ramping activity (Roitman & Shadlen, 2002; Steinemann et al., 2024). Notably, the population signal was predictive of choice and wager during identical epochs (Figure 5B & C), consistent with a parallel process. An interesting anomaly was the temporal offset between choice and confidence decoders at the end of the trial (Figure 5B,C). This offset was not driven by a lag between the trial-by-trial predictions during decision formation, but such a lag did emerge around saccade onset (Figure 6B). This could be a correlate of top-

476 down signals related to attention, expectation, or affect—but another possibility is that it reflects
477 reevaluation of the evidence for guiding subsequent decisions (van den Berg et al., 2016b; Lak et
478 al., 2020 Zylberberg, 2021). One could test this by manipulating neural activity during the peri-
479 decision epoch in a task where optimal behavior depends on confidence in earlier choices. Lastly,
480 we found that the strength of the decoder's choice prediction ('certainty') was highly correlated with
481 the probability of a high bet established by the wager decoder (Figure 6A). To our knowledge this
482 is the first study to directly link a neural representation of choice certainty with a neural prediction
483 of confidence as measured behaviorally.
484
485 **Caveats and conclusions**
486 The main limitation of this study is that confidence is mapped onto a stable motor action, namely
487 the saccade to a high or low target whose positions did not change within a session. We found this
488 to be necessary for achieving consistent behavioral performance, but it does present a challenge
489 for disentangling cognitive signals from those involved in motor planning. Could our results be
490 explained as merely the concurrent evolution of multiple motor plans? We do not think so. First, the
491 signatures of noisy evidence accumulation (Figure 4E, Figure 5A) suggest that ramping activity in
492 LIP reflects more than simple motor planning. Second, although there is evidence for parallel
493 computation of competing plans (Cisek & Kalaska, 2002), even a motor-centric interpretation of our
494 findings would be something different: simultaneous encoding of two dimensions of a motor plan
495 (horizontal and vertical), corresponding to distinct transformations of the input (a categorical
496 judgment vs. the quality of the evidence and/or decision process). It was not a foregone conclusion
497 that the choice and wager dimensions could be computed in parallel; in fact, a 2-D decision with
498 similar task structure (simultaneous report of color and motion) exposed a bottleneck preventing
499 parallel incorporation of multiple evidence streams into a single DV (Kang et al., 2021; Jeurissen,
500 Zylberberg, & Shadlen, unpublished observations).
501 Importantly, the current findings are not incompatible with post-decision processing for
502 confidence. Indeed, even though the peri-decision time window is highly compressed in our task,

there are hints of continued processing related to confidence and potential CoMs (Figure 2C; Figure 5B). Investigating this window further could be useful for testing mechanisms of confidence readout, which for a 2D accumulator requires querying the state of the losing race as well as estimating elapsed time (Kiani et al., 2014a). The contributions to this readout mechanism from sensory populations versus top-down signals remain to be teased apart. Simultaneous recordings across multiple brain areas will be essential for resolving this question, bringing us closer to understanding how neural dynamics convert sensation into belief.

Acknowledgements

We thank members of the lab for their insight and discussions. Additionally, we are grateful to Ofelia Garalde for helping with animal assistance, and Justin Killebrew, Bill Nash, and Bill Quinlan for technical assistance. This work was supported by the National Institute of Neurological Disorders and Stroke (RF1NS132910), the E. Matilda Ziegler Foundation for the Blind, and a Whitehall Foundation research grant. CF is supported by the France-Merrick Foundation.

REFERENCES

Adelson, E. H., & Bergen, J. R. (1985). Spatiotemporal energy models for the perception of motion. *Journal of the Optical Society of America A*, *2*(2), 284. https://doi.org/10.1364/JOSAA.2.000284

Adler, W. T., & Ma, W. J. (2018). Comparing Bayesian and non-Bayesian accounts of human confidence reports. *PLoS Computational Biology*, *14*(11), 1–34. https://doi.org/10.1371/journal.pcbi.1006572

Audley, R. J. (1960). A stochastic model for individual choice behavior. Psychological Review 67, 1.

Balsdon, T., Mamassian, P., & Wyart, V. (2021). Separable neural signatures of confidence during perceptual decisions. *BioRxiv*, 2021.04.08.439033. https://www.biorxiv.org/content/10.1101/2021.04.08.439033v2%0Ahttps://www.biorxiv.org/content/10.1101/2021.04.08.439033v2.abstract

Bang, D., Kishida, K. T., Lohrenz, T., White, J. P., Laxton, A. W., Tatter, S. B., Fleming, S. M., & Montague, P. R. (2020). Sub-second Dopamine and Serotonin Signaling in Human Striatum during Perceptual Decision-Making. *Neuron*, 1–12. https://doi.org/10.1016/j.neuron.2020.09.015

Baranski, J. V., & Petrusic, W. M. (1998). Probing the Locus of Confidence Judgments: Experiments on the Time to Determine Confidence. *Journal of Experimental Psychology: Human Perception and Performance*, *24*(3), 929–945. https://doi.org/10.1037/0096-1523.24.3.929

Bollimunta, A., Totten, D., & Ditterich, J. (2012). Neural dynamics of choice: Single-trial analysis of decision-related activity in parietal cortex. *Journal of Neuroscience*, *32*(37), 12684–12701. https://doi.org/10.1523/JNEUROSCI.5752-11.2012

Boundy-Singer, Z. M., Ziemba, C. M., & Goris, R. L. T. (2022). Confidence reflects a noisy decision reliability estimate. *Nature Human Behaviour*, *7*(1), 142–154. https://doi.org/10.1038/s41562-022-01464-x

Boundy-Singer, Z. M., Ziemba, C. M., & Goris, R. L. T. (2024a). How does V1 population activity inform pereceptual certainty? *Journal of Vision.* 24(6)12: 1-17.

Boundy-Singer, Z. M., Ziemba, C. M., & Goris, R. L. T. (2024b). Sensory population activity reveals confidence computations in the primate visual system. *bioRxiv.*

Caspi, A., Beutter, B. R., & Eckstein, M. P. (2004). The time course of visual information accrual guiding eye movement decisions during visual search. *Proceedings of the National Academy of Sciences*, *4*(8), 743. http://jov.arvojournals.org/Article.aspx?doi=10.1167/4.8.743%5Cnpapers3://publication/doi/10.1167/4.8.743

Chen, X., Zirnsak, M., Vega, G. M., Govil, E., Lomber, S. G., & Moore, T. (2020). Parietal Cortex Regulates Visual Salience and Salience-Driven Behavior. *Neuron*, *106*(1), 177-187.e4. https://doi.org/10.1016/j.neuron.2020.01.016

Churchland, A. K., Kiani, R., R., C., Wang, X.-J., Pouget, A., & Shadlen, M. N. (2011). Variance as a signature of neural computations during decision-making. *Neuron*, *69*(2), 818–831. https://doi.org/10.1038/mp.2011.182.doi

Cisek, P., & Kalaska, J. F. (2002). Simultaneous encoding of multiple potential reach directions in dorsal premotor cortex. *Journal of Neurophysiology*, *87*(2), 1149–1154. https://doi.org/10.1152/jn.00443.2001

Daniel, R., & Pollmann, S. (2012). Striatal activations signal prediction errors on confidence in the absence of external feedback. *NeuroImage*, *59*(4), 3457–3467. https://doi.org/10.1016/j.neuroimage.2011.11.058

David Smith, J., Couchman, J. J., & Beran, M. J. (2012). The highs and lows of theoretical interpretation in animal-metacognition research. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *367*(1594), 1297–1309. https://doi.org/10.1098/rstb.2011.0366

de Lafuente, V., Jazayeri, M., & Shadlen, M. N. (2015). Representation of accumulating evidence for a decision in two parietal areas. *Journal of Neuroscience*, *35*(10), 4306–4318. https://doi.org/10.1523/JNEUROSCI.2451-14.2015

Desender, K., Donner, T. H., & Verguts, T. (2021). Dynamic expressions of confidence within an evidence accumulation framework. *Cognition*, *207*(May 2020), 104522. https://doi.org/10.1016/j.cognition.2020.104522

Desender, K., Ridderinkhof, K. R., & Murphy, P. R. (2021). Understanding neural signals of post-decisional performance monitoring: An integrative review. *ELife*, *10*, 1–21. https://doi.org/10.7554/eLife.67556

Dotan, D., Meyniel, F., & Dehaene, S. (2018). On-line confidence monitoring during decision making. *Cognition*, *171*(February 2017), 112–121. https://doi.org/10.1016/j.cognition.2017.11.001

Dou, W., Martinez Arango, L. J., Castaneda, O. G., Arellano, L., Mcintyre, E., Yballa, C., & Samaha, J. (2024). Neural Signatures of Evidence Accumulation Encode Subjective Perceptual Confidence Independent of Performance. *Psychological Science*. https://doi.org/10.1177/09567976241246561

Double, K. S., & Birney, D. P. (2024). Confidence judgments interfere with perceptual decision making. *Scientific Reports*, *14*(1), 1–10. https://doi.org/10.1038/s41598-024-64575-7

Drugowitsch, J., Mendonça, A. G., Mainen, Z. F., & Pouget, A. (2019). Learning optimal decisions with confidence. *Proceedings of the National Academy of Sciences of the United States of America*, *116*(49), 24872–24880. https://doi.org/10.1073/pnas.1906787116

Drugowitsch, J., Moreno-Bote, R., Churchland, A. K., Shadlen, M. N., & Pouget, A. (2012). The Cost of Accumulating Evidence in Perceptual Decision Making. *Journal of Neuroscience*, *32*(11), 3612–3628. https://doi.org/10.1523/JNEUROSCI.4010-11.2012

Fan, Y., Doi, T., Gold, J. I., & Ding, L. (2024). Neural Representations of Post-Decision Accuracy and Reward Expectation in the Caudate Nucleus and Frontal Eye Field. *Journal of Neuroscience*, *44*(2). https://doi.org/10.1523/JNEUROSCI.0902-23.2023

Fetsch, C. R., Kiani, R., Newsome, W. T., & Shadlen, M. N. (2014). Effects of Cortical Microstimulation on Confidence in a Perceptual Decision. *Neuron*, *83*(4), 797–804. https://doi.org/10.1016/j.neuron.2014.07.011

Fetsch, C. R., Kiani, R., & Shadlen, M. N. (2014). Predicting the accuracy of a decision: A neural mechanism of confidence. *Cold Spring Harbor Symposia on Quantitative Biology*, *79*, 185–197. https://doi.org/10.1101/sqb.2014.79.024893

Gherman, S., & Philiastides, M. G. (2018). Human VMPFC encodes early signatures of confidence in perceptual decisions. *ELife*, *7*, 1–28. https://doi.org/10.7554/eLife.38293

Gnadt, J. W., & Andersen, R. A. (1988). Memory related motor planning activity in posterior parietal cortex of macaque. *Experimental Brain Research*, *70*(1), 216–220.

Guggenmos, M., Wilbertz, G., Hebart, M. N., & Sterzer, P. (2016). Mesolimbic confidence signals guide perceptual learning in the absence of external feedback. *ELife*, *5*(MARCH2016), 1–19. https://doi.org/10.7554/eLife.13388

Hanks, T. D., Ditterich, J., & Shadlen, M. N. (2006). Microstimulation of macaque area LIP affects decision-making in a motion discrimination task. *Nature Neuroscience*, *9*(5), 682–689. https://doi.org/10.1038/nn1683

Hanks, T. D., Mazurek, M. E., Kiani, R., Hopp, E., & Shadlen, M. N. (2011). Elapsed Decision Time Affects the Weighting of Prior Probability in a Perceptual Decision Task. *Journal of Neuroscience*, *31*(17), 6339–6352. https://doi.org/10.1523/JNEUROSCI.5613-10.2011

Herregods, S., Denmat, P. Le, & Desender, K. (2023). Modelling Speed-Accuracy Tradeoffs in the Stopping Rule for Confidence Judgments. *BioRxiv*, 1–58.

Hilgenstock, R., Weiss, T., & Witte, O. W. (2014). You'd Better Think Twice: Post-Decision Perceptual Confidence. *NeuroImage*, *99*, 323–331. https://doi.org/10.1016/j.neuroimage.2014.05.049

Insabato, A., Pannunzi, M., Rolls, E. T., & Deco, G. (2010). Confidence-related decision making. *Journal of Neurophysiology*, *104*(1), 539–547. https://doi.org/10.1152/jn.01068.2009

Jeurissen, D., Shushruth, S., El-Shamayleh, Y., Horwitz, G. D., & Shadlen, M. N. (2022). Deficits in decision-making induced by parietal cortex inactivation are compensated at two timescales. *Neuron*, *110*(12), 1924-1931.e5. https://doi.org/10.1016/j.neuron.2022.03.022

Kang, Y. H. R., Löffler, A., Jeurissen, D., Zylberberg, A., Wolpert, D. M., & Shadlen, M. N. (2021). Multiple decisions about one object involve parallel sensory acquisition but time-multiplexed evidence incorporation. *ELife*, *10:e63721*. https://doi.org/10.7554/eLife.63721

Katz, L. N., Yates, J. L., Pillow, J. W., & Huk, A. C. (2016). Dissociated functional significance of decision-related activity in the primate dorsal stream. *Nature*, *535*(7611), 285–288. https://doi.org/10.1038/nature18617

Kaufman, M. T., Churchland, M. M., Ryu, S. I., & Shenoy, K. V. (2015). Vacillation, indecision and hesitation in moment-by-moment decoding of monkey motor cortex. *ELife*, *4*:e04677. https://doi.org/10.7554/eLife.04677

Kellogg, W. N. (1931). The Time of Judgment in Psychometric Measures. *The American Journal of Psychology*, *43*(1), 65–86.

Kepecs, A., & Mainen, Z. F. (2012). A computational framework for the study of confidence in humans and animals. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *367*(1594), 1322–1337. https://doi.org/10.1098/rstb.2012.0037

Kepecs, A., Uchida, N., Zariwala, H. A., & Mainen, Z. F. (2008). Neural correlates, computation and behavioural impact of decision confidence. *Nature*, *455*(7210), 227–231. https://doi.org/10.1038/nature07200

Khalvati, K., Kiani, R., & Rao, R. P. N. (2021). Bayesian inference with incomplete knowledge explains perceptual confidence and its deviations from accuracy. *Nature Communications*, *12*(1), 1–16. https://doi.org/10.1038/s41467-021-25419-4

Kiani, R., Corthell, L., & Shadlen, M. N. (2014a). Choice certainty is informed by both evidence and decision time. *Neuron*, *84*(6), 1329–1342. https://doi.org/10.1016/j.neuron.2014.12.015

Kiani, R., Cueva, C. J., Reppas, J. B., & Newsome, W. T. (2014b). Dynamics of neural population responses in prefrontal cortex indicate changes of mind on single trials. *Current Biology*, *24*(13), 1542–1547. https://doi.org/10.1016/j.cub.2014.05.049

Kiani, R., Hanks, T. D., & Shadlen, M. N. (2008). Bounded Integration in Parietal Cortex Underlies Decisions Even When Viewing Duration Is Dictated by the Environment. *Journal of Neuroscience*, *28*(12), 3017–3029. https://doi.org/10.1523/JNEUROSCI.4761-07.2008

Kiani, R., & Shadlen, M. N. (2009). Representation of confidence associated with a decision by neurons in the parietal cortex. *Science*, *324*(5928), 759–764. https://doi.org/10.1126/science.1169405

Komura, Y., Nikkuni, A., Hirashima, N., Uetake, T., & Miyamoto, A. (2013). Responses of pulvinar neurons reflect a subject's confidence in visual categorization. *Nature Neuroscience*, *16*(6), 749–755. https://doi.org/10.1038/nn.3393

Lak, A., Costa, G. M., Romberg, E., Koulakov, A. A., Mainen, Z. F., & Kepecs, A. (2014). Orbitofrontal cortex is required for optimal waiting based on decision confidence. *Neuron*, *84*(1), 190–201. https://doi.org/10.1016/j.neuron.2014.08.039

Lak, A., Hueske, E., Hirokawa, J., Masset, P., Ott, T., Urai, A. E., Donner, T. H., Carandini, M., Tonegawa, S., Uchida, N., & Kepecs, A. (2020). Reinforcement biases subsequent perceptual decisions when confidence is low: A widespread behavioral phenomenon. *ELife*, *9*:e49834. https://doi.org/10.7554/eLife.49834

Latimer, K. W., Yates, J. L., Huk, M. L. R. M. A. C., & Pillow, J. W. (2015). Single-trial spike trains in parietal cortex reveal discrete steps during decision-making. *Science*, 34(6244), 184–188. https://doi.org/10.1093/annonc/mdx490/4098869/Hybrid-capture-based-genomic-profiling-of

Lewis, J. W., & Essen, D. C. V. A. N. (2000). Corticocortical Connections of Visual , Sensorimotor , and Multimodal Processing Areas in the Parietal Lobe of the Macaque Monkey. *The Journal of Comparative Neurology*, 428(1), 112–137.

Li, X., Su, R., Chen, Y., & Yang, T. (2023). Optimal policy for uncertainty estimation concurrent with decision making. *Cell Reports*, *42*(3), 112232. https://doi.org/10.1016/j.celrep.2023.112232

Lisi, M., Mongillo, G., Milne, G., Dekker, T., & Gorea, A. (2021). Discrete confidence levels revealed by sequential decisions. *Nature Human Behaviour*, *5*(2), 273–280. https://doi.org/10.1038/s41562-020-00953-1

Ma, W. J., & Jazayeri, M. (2014). Neural Coding of Uncertainty and Probability. *Annual Review of Neuroscience*, *37*(1), 205–220. https://doi.org/10.1146/annurev-neuro-071013-014017

McPeek, R. M., Skavenski, A. A., & Nakayama, K. (2000). Concurrent processing of saccades in visual search. *Vision Research*, *40*(18), 2499–2516. https://doi.org/10.1016/S0042-6989(00)00102-4

Meirhaeghe, N., Riehle, A., & Brochier, T. (2023). Parallel movement planning is achieved via an optimal preparatory state in motor cortex. *Cell Reports*, *42*(2), 112136. https://doi.org/10.1016/j.celrep.2023.112136

Meyniel, F., Sigman, M., & Mainen, Z. F. (2015). Confidence as Bayesian Probability: From Neural Origins to Behavior. *Neuron*, *88*(1), 78–92. https://doi.org/10.1016/j.neuron.2015.09.039

Middlebrooks, P. G., & Sommer, M. A. (2011). Metacognition in monkeys during an oculomotor task. *Journal of Experimental Psychology: Learning Memory and Cognition*, *37*(2), 325–337. https://doi.org/10.1037/a0021611

Moran, R., Teodorescu, A. R., & Usher, M. (2015). Post choice information integration as a causal determinant of confidence: Novel data and a computational account. *Cognitive Psychology*, *78*, 99–147. https://doi.org/10.1016/j.cogpsych.2015.01.002

Murphy, P. R., Robertson, I. H., Harty, S., & O'Connell, R. G. (2015). Neural evidence accumulation persists after choice to inform metacognitive judgments. *ELife*, *4*:e11946. https://doi.org/10.7554/eLife.11946

Nguyenkim, J. D., & DeAngelis, G. C. (2003). Disparity-based coding of three-dimensional surface orientation by macaque middle temporal neurons. *Journal of Neuroscience*, *23*(18), 7117–7128. https://doi.org/10.1523/jneurosci.23-18-07117.2003

Nienborg, H., & Cumming, B. G. (2009). Decision-related activity in sensory neurons reflects more than a neurons causal effect. *Nature*, *459*(7243), 89–92. https://doi.org/10.1038/nature07821

O'Connell, R. G. & Kelly, S. P. Neurophysiology of Human Perceptual Decision-Making. Annu Rev Neurosci 44, 495–516 (2021).

Orozco, S. P., Albert, S. T., & Shadmehr, R. (2021). Adaptive control of movement deceleration during saccades. *PLoS Computational Biology*, *17*(7), 1–30. https://doi.org/10.1371/journal.pcbi.1009176

Pachitariu, M., Steinmetz, N. A., Kadir, S. N., Carandini, M. & Harris, K. D. (2016). Fast and accurate spike sorting of high-channel count probes with KiloSort. Adv. Neural Infor. Process. Syst. 29, 4455–4464.

Peixoto, D., Verhein, J. R., Kiani, R., Kao, J. C., Nuyujukian, P., Chandrasekaran, C., Brown, J., Fong, S., Ryu, S. I., Shenoy, K. V., & Newsome, W. T. (2019). Decoding and perturbing decision states in real time. *Nature*. 591, 604-609. https://doi.org/10.1101/681783

Pereira, M., Faivre, N., Iturrate, I., Wirthlin, M., Serafini, L., Martin, S., Desvachez, A., Blanke, O., de Ville, D. Van, & del Millán, J. R. (2020). Disentangling the origins of confidence in speeded perceptual judgments through multimodal imaging. *Proceedings of the National Academy of Sciences of the United States of America*, *117*(15), 8382–8390. https://doi.org/10.1073/pnas.1918335117

Pleskac, T. J., & Busemeyer, J. R. (2010). Two-stage dynamic signal detection: A theory of choice, decision time, and confidence. *Psychological Review*, *117*(3), 864–901. https://doi.org/10.1037/a0019737

Pouget, A., Drugowitsch, J., & Kepecs, A. (2016). Confidence and certainty: Distinct probabilistic quantities for different goals. *Nature Neuroscience*, *19*(3), 366–374. https://doi.org/10.1038/nn.4240

Purcell, B. A., & Kiani, R. (2016). Hierarchical decision processes that operate over distinct timescales underlie choice and changes in strategy. *Proceedings of the National Academy of Sciences*, *113*(31), E4531–E4540. https://doi.org/10.1073/pnas.1524685113

Rescorla RA, and Wagner AR (1972) A theory of Pavlovian conditioning: variations in the effectiveness of reinforcement and nonreinforcement. In: Classical conditioning II: current research and theory (Black AH, Pro-kasy WF, eds.). New York: Appleton-Century-Crofts.

Resulaj, A., Kiani, R., Wolpert, D. M., & Shadlen, M. N. (2009). Changes of mind in decision-making. *Nature*, *461*(7261), 263–266. https://doi.org/10.1038/nature08275

Roitman, J. D., & Shadlen, M. N. (2002). Response of neurons in the lateral intraparietal area during a combined visual discrimination reaction time task. *Journal of Neuroscience*, *22*(21), 9475–9489. https://doi.org/10.1523/jneurosci.22-21-09475.2002

Saleem KS, and Logothetis NK. (2006). *A Combined MRI and Histology Atlas of the Rhesus Monkey Brain*. Amsterdam: Academic Press.

Sanders, J. I., Hangya, B., & Kepecs, A. (2016). Signatures of a Statistical Computation in the Human Sense of Confidence. *Neuron*, *90*(3), 499–506. https://doi.org/10.1016/j.neuron.2016.03.025

Sarafyazd, M., & Jazayeri, M. (2019). Hierarchical reasoning by neural circuits in the frontal cortex. *Science*, *364*(6441). https://doi.org/10.1126/science.aav8911

Shadlen, M. N., & Kiani, R. (2013). Decision making as a window on cognition. *Neuron*, *80*(3), 791–806. https://doi.org/10.1016/j.neuron.2013.10.047

Shadmehr, R., Reppert, T. R., Summerside, E. M., Yoon, T., & Ahmed, A. A. (2019). Movement Vigor as a Reflection of Subjective Economic Utility. *Trends in Neurosciences*, *42*(5), 323–336. https://doi.org/10.1016/j.tins.2019.02.003

Shan, H., Moreno-Bote, R., & Drugowitsch, J. (2019). Family of closed-form solutions for two-dimensional correlated diffusion processes. *Physical Review E*, *100*(3), 321-332. https://doi.org/10.1103/PhysRevE.100.032132

Shekhar, M. & Rahnev, D. How Do Humans Give Confidence? A Comprehensive Comparison of Process Models of Perceptual Metacognition. J. Exp. Psychol.: Gen. 153, 656–688 (2023).

Shushruth, X. S., Mazurek, M., & Shadlen, M. N. (2018). Comparison of Decision-Related Signals in Sensory and Motor Preparatory Responses of Neurons in Area LIP. *Journal of Neuroscience*, *38*(28), 6350–6365. https://doi.org/10.1523/JNEUROSCI.0668-18.2018

Siegle, J. H. et al. (2017). Open Ephys: an open-source, plugin-based platform for multichannel electrophysiology. *J. Neural Eng.* **14**, 045003.

Steinemann, N. A., Stine, G. M., Trautmann, E. M., Zylberberg, A., Wolpert, D. M., & Shadlen, M. N. (2023). Direct observation of the neural computations underlying a single decision. *eLife.* 12:RP90859.

Stolyarova, A., Rakhshan, M., Hart, E. E., O'Dell, T. J., Peters, M. A. K., Lau, H., Soltani, A., & Izquierdo, A. (2019). Contributions of anterior cingulate cortex and basolateral amygdala to

decision confidence and learning under uncertainty. *Nature Communications*, *10*(1). https://doi.org/10.1038/s41467-019-12725-1

Stringer, C., Pachitariu, M., Steinmetz, N. *et al.* High-dimensional geometry of population responses in visual cortex. *Nature* **571**, 361–365 (2019). https://doi.org/10.1038/s41586-019-1346-5

Sutton RS, Barto AG (1998) Reinforcement learning: an introduction. Cambridge, MA: MIT

van den Berg, R., Anandalingam, K., Zylberberg, A., Kiani, R., Shadlen, M. N., & Wolpert, D. M. (2016a). A common mechanism underlies changes of mind about decisions and confidence. *ELife*, *5:e12192*. https://doi.org/10.7554/eLife.12192

van den Berg, R., Zylberberg, A., Kiani, R., Shadlen, M. N., & Wolpert, D. M. (2016b). Confidence Is the Bridge between Multi-stage Decisions. *Current Biology*, *26*(23), 3157–3168. https://doi.org/10.1016/j.cub.2016.10.021

Wang, S., Falcone, R., Richmond, B., & Averbeck, B. B. (2023). Attractor dynamics reflect decision confidence in macaque prefrontal cortex. *Nature Neuroscience*, *26*(11), 1970–1980. https://doi.org/10.1038/s41593-023-01445-x

Xue, K., Zheng, Y., Rafiei, F., & Rahnev, D. (2023). The timing of confidence computations in human prefrontal cortex. *Cortex*, *168*, 167–175. https://doi.org/10.1016/j.cortex.2023.08.009

Zhou, Y., & Freedman, D. J. (2019). Posterior parietal cortex plays a causal role in perceptual and categorical decisions. *Science*. 365(6449), 180-185. https://doi.org/10.1007/978-1-4614-7320-6_310-1

Zylberberg, A. (2021). Decision prioritization and causal reasoning in decision hierarchies. *PLoS Computational Biology*. 17(12):e1009688. https://doi.org/10.1371/journal.pcbi.1009688

Zylberberg, A., Barttfeld, P., Sigman, M., & Pereira, A. (2012). The construction of confidence in a perceptual decision. *Front. Integr. Neurosci. 6*(79), 1–10. https://doi.org/10.3389/fnint.2012.00079

METHODS

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Two male rhesus monkeys (*Macaca mulatta,* 6-8 years old, 8-10 kilograms) were kept and handled according to National Institutes of Health guidelines and the Institutional Animal Care and Use Committee at Johns Hopkins University. Standard sterile surgical procedures were performed to place a PEEK recording chamber (Rogue Research) and titanium head post under isoflurane anesthesia in a dedicated operating suite. The recording chamber was positioned over a craniotomy above the right posterior parietal cortex of both animals for access to the intraparietal sulcus and posterior third of the superior temporal sulcus. The chamber and head post were held in place using dental acrylic anchored with ceramic bone screws.

METHOD DETAILS

Experimental apparatus

Monkeys were seated in a custom-built primate chair in a sound-insulated booth facing a visual display (ViewPixx, VPixx Technologies; resolution 1080x960, refresh rate 120 Hz; viewing distance 52 cm) and infrared video eye tracker (Eyelink 1000 Plus, SR Research). Experiments were controlled by a Linux PC running a modified version of the PLDAPS system (Version 4.1, Eastman & Huk, 2012) in MATLAB (The MathWorks). Visual stimuli were generated using Psychophysics Toolbox 3.0 (Brainard, 1997). For correct responses, the monkey was given a fluid reward that was dispensed using a solenoid-gated system.

Neurophysiology

Recording probes (32- or 128-channel Deep Array, Diagnostic Biochips) were positioned with the aid of a PEEK grid secured inside the recording chamber. A sharpened guide tube was inserted through a grid hole so that the tip of the tube just punctured the dura, then a probe was advanced through the guide tube into the brain using a motorized microdrive (40mm MEM Drive, Thomas

830 Recording). Bandpass-filtered voltage signals were collected using the Open Ephys acquisition
831 board and software (Siegle et al., 2017). Post-hoc analysis for identifying single neurons and multi-
832 unit clusters was done using Kilosort 2.0 (Pachitariu et al., 2016; Stringer et al., 2019). Further
833 curation was done using the phy2 software (https://phy.readthedocs.io/en/latest/). Data analysis
834 was performed with custom MATLAB code.

835 Targeting of LIPv was achieved by selection of grid locations based on a post-surgical
836 structural MRI scan in which the chamber and grid holes were well visualized. We compared the
837 MR images to published reports and atlases (Lewis and Van Essen, 2000; Saleem & Logothetis,
838 2006) to estimate the depth of LIPv (typically 8-12 mm from the dura in our vertical penetration
839 angle) and corroborated the targeting using white-gray matter transitions and physiological
840 response properties during the mapping tasks described below. After reaching the target, we let
841 the probe settle for 30-60 minutes before the start of the experiment. A total of 407 neurons (single
842 = 148 & multi = 56 in monkey H, single = 107 & multi = 93 G) were collected over 29 sessions (12
843 for monkey H, 17 for monkey G).

844

845 **Memory saccade task**

846 Sessions began with a standard memory-guided saccade task to identify neurons with
847 spatially selective activity during the delay period (Gnadt & Andersen, 1998) and to coarsely map
848 their response fields (RFs). Monkeys were instructed to gaze at a central fixation point (1.5° radius
849 acceptance window), after which a red target (0.42° diameter circle) was flashed for 100 ms,
850 located in one of several locations evenly spaced in polar coordinates. The coordinates consisted
851 of 3 different radii (eccentricities) and 10 or 12 angular positions, giving a total of 30 or 36 unique
852 target locations. Each target location was presented 10 times in pseudorandom order, requiring a
853 total of 300 or 360 trials. While fixating, the monkey had to remember the location of the target, and
854 after a delay of 0.8 s the fixation point was extinguished, instructing the monkey to make a saccade
855 to the remembered location. RFs were estimated online during/after the memory saccade block by
856 acquiring multi-unit spikes (threshold crossings) on each recording channel and plotting the mean

857 firing rate during the memory delay as a function of target location in a 2D heat map. These RF
858 maps guided the placement of the four targets for the main decision task, such that one or two of
859 the targets overlapped the RF of multiple neurons in the recorded ensemble.
860
861 **Main task**
862 The monkeys were trained to perform a reaction-time direction discrimination task with
863 simultaneous report of choice and confidence ('peri-decision wagering'; Figure 1A). To initiate a
864 trial the animals acquired fixation on a target on the center of the screen (0.21° diameter). After a
865 delay of 0.5 s four targets appeared, positioned diagonally from the center of the screen, each
866 representing a choice (left or right) and a wager, or bet (high or low). The targets representing high
867 bets were always placed in the upper-left and upper-right quadrants while the low-bet targets were
868 placed in the lower quadrants (left high: x = -7.3±2.1°, y = 7.3±1.5°; left low: x = -6.9±2.3°, y = -
869 3.4±1.4°; right high: x = 7.3±2.1°, y = 7.3±1.5°; right low: x = 6.9±2.3°, y = -3.4±1.4°). Each left-right
870 pair was presented symmetrically around the vertical meridian, but high-bet targets were typically
871 2-5° further from the fixation point than low-bet targets, in order to counteract the monkeys'
872 tendency to bet high more often than low.
873 After another brief delay (0.3-0.6 s, truncated exponential), a dynamic random-dot motion
874 (RDM) stimulus was presented in a circular aperture. Motion strength, or coherence, was sampled
875 uniformly on each trial from the set {0%, ±3.2%, ±6.4%, ±12.8%, ±25.6%, ±51.2%}, where positive
876 is rightward and negative is leftward. The stimulus was constructed as three independent sets of
877 dots (Roitman & Shadlen, 2002), each appearing for a given video frame then reappearing three
878 frames (25 ms) later. Upon reappearing, a given dot was either repositioned horizontally to
879 generate apparent motion in the assigned direction (speed = 2-16°/s, held constant within a
880 session) with probability given by the coherence on that trial, or otherwise was replotted randomly
881 within the aperture.
882 When ready with a decision, the animal could report its choice and wager by making a
883 single saccade to one of the four targets. When the eyes moved 1.5° away from the target the

RDM and fixation point were extinguished while the four targets remained visible. When the eye position reached one of the four targets, it was required to hold fixation within 1.5° of the target for 0.1 s to confirm the outcome. Lastly, the animal was either rewarded or given a time penalty depending on the conjunction of accuracy (choice corresponding to the sign of coherence) and wager (Figure 1A, right). The penalty for a high-bet error was applied to the subsequent trial where the animal was required to fixate the central target for a longer period of time (2-3 s) prior to RDM onset. Reward sizes (~.21 ml for high and ~.19 ml for low bets) and penalty times were chosen to encourage a wide range of wager probability across different levels of motion strength.

QUANTIFICATION AND STATISTICAL ANALYSIS

Cell selection

To quantify spatial selectivity, we modified a previously described discrimination index (Nguyenkim & DeAngelis, 2003).

$$DDI = \frac{R_{max} - R_{min}}{R_{max} - R_{min} + 2\sqrt{\frac{SSE}{N-M}}}$$

R_{max} and R_{min} are the mean firing rates during the delay period at the target spatial location with the highest and lowest response, respectively. SSE is the sum-squared error around the mean responses, N is the total number of trials, and M is the total number of unique spatial target locations. For all analyses except the population logistic decoder, only units (single- and multi-) with a discrimination index higher than an arbitrary cutoff of 0.45 were kept for further analysis.

Full quantification of RFs was done by fitting a 2D Gaussian to the firing rates during the delay period of the memory saccade task for each target location:

$$FR = A * \frac{1}{2\pi|\Sigma|^{\frac{1}{2}}} \exp\left(-\frac{1}{2}(x-\mu)\Sigma^{-1}(x-\mu)^T\right)$$

where FR is firing rate during the delay period and X is a 2-length vector that contains the X and Y positions of the target locations (in degrees). The fitted variables include $\boldsymbol{A}$ for amplitude, $\boldsymbol{\mu}$ for the 2D mean location of the RDM (X and Y in degrees), and $\boldsymbol{\Sigma}$ which is the covariance matrix of the Gaussian (2x2). To reduce the variables for fitting we set the covariance to 0 and only fit for the

two variances. We then normalized the 2D Gaussian to convert it to a probability density and calculated the final preferred target based on the probabilities corresponding to the 4 targets in the peri-dw task.

For the population logistic decoder, we included not only units with delay period activity in the memory saccade task but accepted any that were well defined single neurons or multi-unit clusters based on careful spike sorting and manual curation. Previous studies have suggested that decision-related activity can be present even in neurons without memory delay-period activity (Meister et al., 2013), and we aimed to maximize the sample size to facilitate single-trial decoding. This broader criterion increased the average number of units per session from 7.5 to 17.3 for monkey H and from 6.2 to 11.2 for monkey G.

Behavioral data analysis

We applied a logistic regression model to fit the proportion of rightward choices, as follows:

$$P_{right} = \frac{1}{1+\exp(\beta_0+\beta_1 Coh+ \beta_2 Wager+\beta_3(Coh)(Wager))}$$

where P_{right} is the probability of a rightward choice, *Coh* is signed motion coherence, *Wager* is the monkey's bet (high/low), β_0 is the overall bias, β_1 estimates the effect of signed motion coherence on choices, β_2 is the weight for the bet, and β_3 is the interaction term (used to calculate whether choice significantly depends on wager). The fitting was done by finding the minimum negative log-likelihood value under a binomial distribution, using *fminsearch* in MATLAB with the Nelder-Mead method. For choices conditioned on high and low bets, we fit two separate logistic regressions after removing the terms to the right of $\beta_1 Coh$ (Figure S1).

The fits to the average reaction times (RTs) as a function of motion strength were done using a Gaussian distribution, defined as follows:

$$RT = A * \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{coh-0}{\sigma}\right)^2\right) + b$$

where A is an amplitude term, coh is the signed motion coherence, σ is the standard deviation controlling the width of the Gaussian, and b is a bias term capturing the fastest mean RT. We fit

this pseudo-gaussian by minimizing root-mean-square error (RMSE) instead of maximizing the likelihood, because of the amplitude value.

To examine the relationship between accuracy and RT, as well as wager and RT, we calculated proportion correct and proportion of high bets grouped by RT, using non-overlapping 100 ms time bins starting 100 ms after motion onset. To test for significance as to whether the trend was decreasing (during the time depicted in Figure 1D) we used a Cochran-Armitage test for trend:

$$T = \sum_{i=1}^{k} t_i (N_{1i} R_2 - N_{2i} R_1)$$
$$Var(T) = \frac{R_1 R_2}{N} \left(\sum_{i=1}^{k} t_i^2 C_i (N - C_i) - 2 \sum_{i=1}^{k-1} \sum_{j=i+1}^{k} t_i t_j C_i C_j\right.$$

where t_i are the weights depicting the trend (in our case linear, so t = [0, 1, 2, 3, 4, …]), N = total number of trials, N_{ab} is total number of trials for group a (accuracy: correct and error; wager: high or low) and b is total number of trials for timepoint b. R_a represents total number of trials for group a irrespective of time, and C_b is total number of trials at timepoint b irrespective of group. The division of T by $Var(T)$ gives a test statistic that can then be used to compute a p-value.

Motion energy analysis

To estimate the temporal weighting of sensory evidence for choice and confidence, we utilized motion stimulus fluctuations to perform a psychophysical reverse correlation analysis on both choice and wager. We convolved each trial's sequence of dots — a 3-dimensional array with the first two dimensions denoting the X-Y coordinates of the dots' center, and the third dimension spanning the number of frames — with two pairs of quadrature spatiotemporal filters (Adelson & Bergen, 1985). The filters were oriented to account for motion in the direction along the choice axis: 0° motion (rightward) and 180° motion (leftward motion). The convolved quadrature pairs were squared and summed to give the local motion energy for both the leftward and rightward directions. These local motion energies were then collapsed across space (first 2 dimensions) to derive the motion energy provided by the stimulus through time, ME(t). To quantify the net motion energy, we took the difference between both directions of motion energy (right minus left; Adelson

& Bergen, 1985; Kiani et al., 2008). To strictly look at the fluctuations around the mean and mitigate potential effects of coherence, we subtracted from each trial the mean motion energy, conditioned on signed coherence, through time. As most meaningful fluctuations occur at smaller coherences, we also decided to restrict the analysis to 0%, 3.2% and 6.4% coherence trials.

For the choice kernels (Figure 2A) we simply averaged the motion energy profiles for all trials with a right choice (red line) or left choice (blue line). The shaded red and blue area represents the standard error of the mean. To evaluate statistical significance of the choice kernels, we used a t-test comparing the motion energy profiles for right and left choices at each time point, applying a Šidák correction matching the number of time samples. For the confidence kernels (Figure 2B), we instead subtracted the motion energy profiles for high vs. low wager trials conditioned on a given choice (correct trials only) and computed the standard error of the difference (shaded area around traces in bottom of Figure 2). To assess significance, we compared the confidence kernel distribution relative to a value of zero using a t-test with Šidák correction for the number of time samples.

Parallel model

To formalize the hypothesis of parallel deliberation for choice and confidence, we utilized a two-dimensional (2D) bounded accumulator model (Kiani et al. 2014), also known as an anticorrelated race. We adapted a recently developed family of closed-form solutions for a 2D correlated diffusion process (Shan et al., 2019) to facilitate fitting of the parameters. By conceptualizing the diffusion process as a Gaussian distribution originating from the third quadrant on a plane with two absorbing bounds, one can employ the method of images (MoI) to calculate the propagation of the probability density of the diffusing particle, i.e. the solution to the Fokker-Planck equation. The constraint making this numerical solution possible limits the discrete number of anti-correlation values that can be modeled, governed by the number of images:

$$rho = \cos\left(-\frac{\pi}{I}\right)$$

where rho is the correlation value and I is the number of images. We selected $I = 4$ or $rho = .7071$ for consistency with previous studies (Kiani et al., 2014; Van Den Berg et al., 2016a).

Specifically, the MoI yields $P\left(v_{right}, v_{left} \middle| C, t\right)$, describing the probability of the accumulator being in a particular position at time t for coherence C. The probability of making a right choice is given by:

$$P(Right|C) = \frac{\int\int P\left(v_{right} = B, v_{left} \middle| C, t\right) dv_{left} * P(t) dt}{\int\left(\int P\left(v_{right} = B, v_{left} \middle| C, t\right) dv_{left} + \int P\left(v_{right}, v_{left} = B \middle| C, t\right) dv_{right}\right) P(t) dt}$$

where B is the bound for terminating the accumulation process. To acquire the decision time distribution, we calculated the difference in the survival probability as follows:

$$P(DT) = \Delta\left(\iint_{-inf}^{Bound} P\left(v_{right}, v_{left} \middle| C, t\right) dv_{right} dv_{left}\right)$$

thereby providing the change in probability at each survival timestep, quantifying the probability of crossing a bound. To then calculate the probability of reaction time (decision time plus sensory/motor delays, referred to as non-decision time or *nonDT*), we convolved this probability density as follows:

$$P(RT) = \int P_{DT}(t - z) P_{nonDT}(z) dz$$

where $P_{nonDT}(z)$ is modeled as a Gaussian distribution with mean (μ_{nonDT}) and standard deviation (σ_{nonDT}).

To calculate the probability of betting high, we first computed the log-odds of a correct choice as a function of the state of the losing accumulator (Kiani et al., 2014), as follows:

$$Logodds = Log\left(\frac{\int P(v_{incorrect}, t | v_{correct} = B, C, t) P(C) dC}{\int P(v_{correct}, t | v_{incorrect} = B, C, t) P(C) dC}\right)$$

where $v_{incorrect}$ is the incorrect accumulator (not matching the sign of coherence) and $v_{correct}$ is the correct accumulator (matching the sign of coherence). This transformation provides a graded

scale for betting that can be transformed into {high, low} responses by applying a cutoff value (θ) that imposes binary outcomes. To obtain the probability of a high bet we computed:

$$P(high\ wager) = \frac{\iint P\left(v_{right} \middle| v_{left} = B, t, C, Logodds < \theta\right) P(t) + \iint P\left(v_{left} \middle| v_{right} = B, t, C, Logodds < \theta\right) P(t)}{\iint P\left(v_{right} \middle| v_{left} = B, t, C\right) P(t) + \iint P\left(v_{left} \middle| v_{right} = B, t, C\right)\ P(t)}$$

where $Logodds < \theta$ indicates integration over the area in which $Logodds$ is less than the cutoff value.

Lastly, we found the best-fitting parameters to the model by using the joint probability of choice, RT, and wager as follows:

$$P(Right, RT, High) = \prod_{tr=1}^{Trials} P(Right|RT, High) * P(RT|High) * P(High)$$

We calculated the negative log likelihood of this by:

$$-\log\left(P(Right, RT, High)\right) = -\sum_{tr=1}^{Trials} \log\left(P(Right|RT, High)\right) + \log(P(RT|High)) + \log\left(P(High)\right)$$

The full model had 6 or 7 free parameters, differing slightly between the two animals (see table below). These included drift rate (K), bound (B), mean non-decision time (μ_{nonDT}), urgency max (u_M), urgency half-life ($u_{\frac{1}{2}}$), and log-odds cutoff (θ). The urgency signal followed a hyperbolic function for monkey G, requiring both parameters, and a linear function for monkey H, requiring one parameter. Also, monkey H required separate non-decision time means for left and right choices, and a 'wagering offset' (b_w) capturing a tendency to occasionally bet low even for the highest coherence. Non-decision time standard deviation (σ_{nonDT}) was not a free parameter and was instead established using psychophysical kernels (van den Berg et al., 2016a, Methods section). Fitting was done using MATLAB's built-in function *fminsearch* applied in a grid search method with 30 different starting points.

Monkey	Parameters	Number of Parameters

H	K, B, $\mu_{nonDT,Coh+}$, $\mu_{nonDT,Coh-}$, u_M, θ, b_w	7
G	K, B, μ_{nonDT}, u_M, $u_{\frac{1}{2}}$, θ	6

Serial model

The serial model was constructed as a sequence of two 1D bounded accumulator (drift-diffusion) models, one for choice followed by a second for confidence (Figure 1C; Moran et al., 2015; Herregods et al., 2023). There are 6 main parameters: drift rate (K), choice bound (B_c), high wager bound (B_h), low wager bound (B_l), mean non-decision time (μ_{nonDT}), and urgency signal for only the confidence accumulator $(u_{M,Conf})$. Parameter estimation predominantly followed the same logic as the parallel model, using the joint distribution of choice, RT, and wager to fit the data, with a few minor differences. Computing the probability of rightward choice was similar to the parallel model and used the following formula:

$$P(Right|C) = \frac{\int P(v_1 = B_c\,|C,t)P(t)dt}{\int\left(P(v_1 = B_c\,|C,t)+P(v_1 = -B_c\,|C,t)\right)P(t)dt}$$

where v is the accumulator variable for the first 1D accumulator, and the other parameters are similar to the parallel model. The decision time distribution was calculated as:

$$P(DT) = \int P_{choice}(t-z)P_{wager}(z)dz$$

where P_{choice} is the probability of hitting a bound at time t for the first accumulator and P_{wager} is the probability of hitting a bound at time t for the second accumulator. For RT we followed the same procedure as the previous model. Lastly, to calculate the probability of betting high, we used the following equation:

$$P(High\ wager|C) = \frac{P(v_2 = B_h|C, Choice = R)}{\int P(v_2|C, Choice = R)\,dv_2}P(Choice = R)$$

$$+\frac{P(v_2 = B_h|C, Choice = L)}{\int P(v_2|C, Choice = L)\, dv_2} P(Choice = L)$$

Monkey	Parameters	Number of parameters
H	K, B_C, B_h, B_l $\mu_{nonDT,Coh+}$, $\mu_{nonDT,Coh-}$, $u_{M,Conf}$, b_w	8
G	K, B_C, B_h, B_l, μ_{nonDT}, $u_{M,Conf}$	6

Changes of mind (CoMs) and changes of confidence

Raw eye position data (sampled at 1000Hz) were converted to velocity and smoothed by applying a 3rd-order low-pass Butterworth filter with a cutoff frequency of 75Hz (Orozco et al., 2021). Eye position is a 2D vector containing X and Y position in degrees, while eye velocity was defined as 1D vector that combines the velocities for both directions (X and Y). To preprocess the data, we first calculated a stricter time for saccade onset by applying a threshold of 20°/s onto the smoothed velocity data. Subsequently, we centered the eye positions by subtracting the average of the first 5 ms before saccade onset. The initial choice and wager were defined by the location of the saccade in 1 of 4 quadrants of the screen 5 ms after saccade detection. The final choice and wager corresponded to the target at which the eye position settled within a 0.1 s grace period following the initial saccade. This provided an initial and final choice and wager for every trial, allowing for simple analyses like those for Figure 2C,D,E. CoM frequencies (Figure 2C,D) were conditioned on the initial outcome; therefore, the frequencies reflect not the proportion of all trials but rather only trials that initially reached a given outcome (error, correct, low, and high). Lastly, for the rightmost panels of Figure 2E the probabilities are reflective of all trials, hence the values are much smaller than on the other plots.

Analysis of neural data based on preferred and chosen target

Two populations of neurons were established based on RF overlap of either the left-high or left-low target. We computed four normalized firing rate responses for both populations of neurons, where each response corresponds to one of the four chosen targets (Figure 4B & 4C). To combine across neurons, we first detrended the responses of each neuron over the coherences of interest [-6.4%, -3.2%, 0%, 3.2%, 6.4%]. Then the firing rates for each neuron (after smoothing each trial with a 0.1 s exponential filter) were averaged over all trials that met 3 conditions: 1) matched the chosen target of interest, 2) contained a coherence of interest, and 3) had a RT longer than 0.3 s. Example neurons in Figure 4A were not detrended or normalized but otherwise the procedure was the same. The colored bars at the top of Figure 4B & 4C indicate statistical significance based on a one-tailed t-test evaluating whether the response of the chosen target (aligned with the neurons' RF) was higher than either of the other three chosen target responses. Given that the testing is done over multiple timepoints, the one-tailed t-test alpha value was corrected using Šidák's correction.

Autocorrelation signatures of standard vs delayed accumulator dynamics

An accumulation of noisy evidence produces characteristic variance and autocorrelation features that can be estimated from single neurons using procedures laid out by previous work (Churchland et al., 2011; de Lafuente et al., 2015; Shushruth et al., 2018; Steinemann et al., 2022). Applying the law of total variance to a doubly stochastic process, the variance in spike rate for a given time bin is a summation of the variance of the underlying latent rate, termed variance of the conditional expectation (VarCE), and the residual variance expected if the latent rate was constant, known as point process variance (PPV). To calculate VarCE, one must subtract out the PPV from the total measured variance. To do this we made two simple assumptions (Churchland et al., 2011): 1) the observed spiking of a neuron follows a stochastic point process mediated by some rate parameter, and 2) that at each time bin the PPV is proportional to the mean count:

$$Var(E[Y_i|X_i]) = Var(Y_i) - E[Var(Y_i|X_i)]$$
$$VarCE(Y_i) = Var(Y_i) - \varphi E[N_i]$$

where Y_i represents the random variable capturing the neuron's spike count at timepoint i, X_i is the random variable for the latent rate at timepoint i, and φ is a constant that is fitted to maximize how well the observed firing rates match an accumulation of independent identically distributed (iid) random numbers. $E[N_i]$ is the mean spike count at timepoint i. In addition, it follows that the law of total covariance is described using a similar equation:

$$Cov(Y_i, Y_j) = Cov(E[Y_i|X_i], E[Y_j|X_j]) + E[Cov(Y_i, Y_j|X_i, X_j)]$$

The first term on the right-hand side is known as the covariance of conditional expectations (CovCE), which is needed to compute the correlation of conditional expectations (CorCE). The second term is the expectation of conditional covariance, and its diagonal is the PPV. To calculate the CorCE, we made another assumption that when $i \neq j$, the expectation of conditional covariance is zero because the variance from the point process should be independent across timepoints (Churchland et al., 2011; although this may be untrue for adjacent time bins due to their shared interspike interval). This simplification makes it possible to state that the CovCE, for $i \neq j$, is equal to the measured covariance, and the diagonal of the CovCE is then the VarCE. It follows that to calculate the CorCE, one must simply divide the CovCE by the VarCE:

$$CorCE(Y_i, Y_j) = \frac{CovCE(Y_i, Y_j)}{\sqrt{VarCE(Y_i) * VarCE(Y_j)}}$$

where i and j are time points. The best fitting φ is then calculated by comparing the formulated CorCE values onto theoretical or simulated correlation values that underlie the hypothesized generative process (e.g., accumulation).

We tested two theoretical autocorrelation patterns, one pertaining to a standard drift diffusion process and the other to a delayed drift diffusion process (Figure 4D). The standard accumulation of iid random numbers was calculated using the following equation:

$$\rho_{i,j} = \sqrt{\min(i,j) / \max(i,j)}$$

We used 6 different timepoints giving 15 unique combinations ($i = 1{:}6\ \&\ j = 1{:}6,\ i \neq j$). For the delayed accumulation of iid random numbers, we used a simulation that accumulated noisy normalized samples of numbers with mean [0.717, 0, -0.717]. We narrowed the simulation to the first 6 timesteps to compare it to the results from the standard accumulation process. Additionally, the delay component was constructed by uniformly sampling a value between 1-6, indicating when the accumulation would begin. Using 10,000 trials we calculated the autocorrelation of the first 6 timesteps for this simulation providing 15 unique combinations ($i = 1{:}6\ \&\ j = 1{:}6,\ i \neq j$). Additionally, we fit the φ, for both models, according to the following steps: 1) calculate the $E[N_i]$, $Var(Y_i)$, and $Cov(Y_i, Y_j)$ from observed spikes, 2) compute $VarCE(Y_i)$ using an initial value of $\varphi = 1$, 3) calculate $CorCE(Y_i, Y_j)$ under the assumptions of mentioned above, 4) calculate the mean squared error (MSE) between the empirical $CorCE(Y_i, Y_j)$ and the theoretical/simulated autocorrelation values, $\rho_{i,j}$, and 5) iteratively update φ until the MSE between the $CorCE(Y_i, Y_j)$ and $\rho_{i,j}$ reached the global minimum.

We used 6x60-ms time bins spanning from 170 ms after motion onset to 530 ms after motion onset. We applied the analysis on trials with coherences of interest [-6.4%, -3.2%, 0%, 3.2%, 6.4%] and reaction times at least 630 ms to minimize bound effects. To combine across neurons, we calculated the mean response for each time bin of each neuron across all trials and subtracted that from the mean response for each time bin conditioned on the signed coherence. This then gives a matrix of residuals that is of size $[neuron * trials \quad x \quad time\ bins]$, which is then used to calculate a covariance matrix. Next, the VarCE is calculated by substituting the raw variance for the diagonal in the covariance matrix, since the diagonal is the normalized population variance, and φ is initiated at a value of 1. To calculate the empirical correlation values, each entry in the covariance matrix is divided by $\sqrt{VarCE(Y_i) * VarCE(Y_j)}$. Lastly, using a fitting procedure

we compare the Fisher z-transformation of the empirical correlation with a Fisher z-transformation of the ideal correlation and minimize the mean squared error (MSE) of the correlation matrix.

Testing for significant difference between standard and delayed accumulation was done using a leave-one-out (LOO) cross-validation method. The metric used to test the validation was the mean absolute percentage error (MAPE). MAPE allowed for comparison between the two models because both models contained different dependent variables ($\rho_{i,j}$). This method provided 15 different MAPE values, for each model, which were then compared using a one-tailed t-test. The model with the lowest percentage error distribution better captured the underlying autocorrelation structure of the data (Figure 4E).

For Figure S5 we recalculated the VarCE but used 6x100-ms time bins spanning -50 ms before motion onset to 550 ms after motion onset. Here we instead applied the analysis on all trials irrespective of coherence. The CorCE statistical outcomes and preference for standard over delayed accumulation process did not differ when using all the coherences. We quantified the shaded errors bars by using a bootstrap method with 100 resamples.

Population logistic decoder

Data were preprocessed by calculating spike counts in 100 ms time windows, stepping every 20 ms, through the first 600 ms after motion onset, and again separately for spikes aligned 400 ms before and 200 ms after saccade onset. We looked only at trials with RTs 400 ms and longer. We used two L1-regularized logistic decoders, one for choice and one for wager:

$$P_t(Right\ choice\ or\ High\ wager) = -\log(1 + \exp(-(w_0 + w * X)) - \xi\|w\|^2$$

where w represents a vector of weights at time t (vector length matches number of neurons), w_0 is a bias term, and X is the spike count vector for each neuron in the session. The best hyperparameter ξ was found using a 50-value grid search between [0 1000], using 5-fold cross validation. The dataset was divided into a training set (90%) and test set (10%), the latter of which was used to calculate the prediction accuracy (Figure 5B). If P_t was greater than 0.5 this would indicate that the decoder predicted either a right choice (for the choice decoder) or high bet (for the

wager decoder). Values below 0.5 would indicate either a left choice or low bet. The performance (accuracy) was defined based on the monkey's choice and wager at the end of the trial. To compute a 'model decision variable' from neural activity we first take the trials in which the monkey made a choice that aligns with the signed coherence (except for 0% coherence). Afterwards, we simply take the log odds of a particular choice (e.g., $\log\left(\frac{P(right)}{1-P(right)}\right)$ for rightward choices) given the ensemble spike counts up to time *t* on a given trial (Kiani et al., 2014; Peixoto et al., 2021). Importantly, for Figure 5A and 5C we used the log odds irrespective of choice, therefore the results combine $\log\left(\frac{P(right)}{1-P(right)}\right)$ for right choices and $\log\left(\frac{1-P(right)}{P(right)}\right)$ for left choices. Because these log odds for right choice and left choice are predominantly symmetric, this increases statistical power. For Figure 5C the log odds for wager are also combined in the following manner: $\log\left(\frac{P(high)}{1-P(high)}\right)$ for high wagers and $\log\left(\frac{1-P(high)}{P(high)}\right)$ for low wagers.

To test whether the model decision variable (DV, log odds) for the choice decoder contained a linear increase that was significantly dependent on motion strength, we utilized a linear regression:

$$DV = \beta_0 + T * \beta_1 + Coh * \beta_2 + T * Coh * \beta_3$$

where β_0 is a bias term, T is time (20 ms time bins for 200-600 ms after motion onset), Coh is motion coherence level, and $\beta_{1,2,3}$ are the weights accompanying the predictor variables. If the β_3 weight was significantly different from zero ($P < 0.05$), then the modulation of log odds by motion strength was deemed significant. Importantly, to compute this linear regression we only used the mean DV shown in Figure 5A, excluding data cut off at the mean RT for each individual coherence.

As previously mentioned, both logistic decoders provide weights for each neuron through time. Therefore, to test whether our sample of neurons contained a single population that contributes approximately equally to choice and wager, we calculated the correlation between the weight magnitudes (irrespective of sign) as well as the distribution of the difference in weight

magnitude. The weights were preprocessed by taking their absolute value because of our interest in magnitude and not direction. To calculate the correlation, we compute the Pearson's correlation between absolute weights for all our neurons at each time point. We collapsed over time by averaging over the first 600 ms after motion onset and last 400 ms before saccade onset (Figure 5D, red line). We computed significance by randomly permutating the choice and wager weights 1000 times (Figure 5D, blue distribution). Additionally, the choice and wager decoder absolute weights were also subtracted from one another to create a distribution which informs whether there is a single population equally contributing to both choice and wager. Significance was evaluated using the Hartigan's dip test, which tests whether a distribution is unimodal or bimodal.

To determine the trial-by-trial relationship between the choice and wager decoder we first tested whether the wager decoder was predictive of the DV (log odds). Trials where categorized as decoded-low or decoded-high confidence by calculating the mean $P(High)$ for the wager decoder from saccade initiation till 200 ms after. Values above (below) .5 indicated a decoded-high (decoded-low) wager. We strictly looked at trials with 0% coherence to try and remove any effects of coherence on the results. Results for Figure 6A were calculated by then averaging the DV on these decoded-high and decoded-low trials. Significance bars were calculated using a one-tailed t-test with Šidák's correction. Importantly, differences in peak DV around saccade onset (Figure 6A) highlight changes in the ratio $\frac{P(right)}{1-P(right)}$.

The temporal offset between the peak log odds for choice and wager (Figure 5B) might suggest that although deliberation for choice and confidence begins simultaneously, the updating accrues some lag as the trial progresses. To test this, we related the log odds information of the choice decoder to the $P(High)$ from the wager decoder. As shown in Figure 6A, there is a link between the model DV and $P(High)$ that can be exploited to understand whether the temporal updating is in sync or if there is some lag. In this case, the alternative hypothesis, given that the peak of $P(High)$ is later than $P(Right)$, is that updating of the DV for choice precedes the updating for wager. Figure 6B was calculated by using the 3 separate 400-ms time windows described in the

legend. The independent variable, which is $P(right)$ from the choice decoder, was corrected so that values near 0.5 (chance level) were close to zero and values moving away from 0.5 in either direction (better predicting left or right choice) became closer to one. This was done using the following equation:

$$corrected\ P(right) = abs(P(right) - .5) * 2$$

In essence, this remapping of $P(Right)$ is changing the property of choice decoding strength so that it is monotonic or linear. This transformation makes it possible for a linear relationship to exist between corrected $P(Right)$ and $P(high)$. To capture this relationship and its dependency on time lag, we applied a linear regression, using the corrected $P(Right)$ at time t, to the dependent variable $P(High)$ at time $t + \Delta t$, where Δt range from +/- 200 ms. To quantify how informative corrected $P(Right)$ is of $P(High)$ we used an R^2 measurement at each time lag. Figure 6B displays a corrected R^2 which is simply the R^2 value after subtracting out the average R^2 values when the time series of the decoders for each trial are randomly permuted.

SUPPLEMENTARY FIGURES

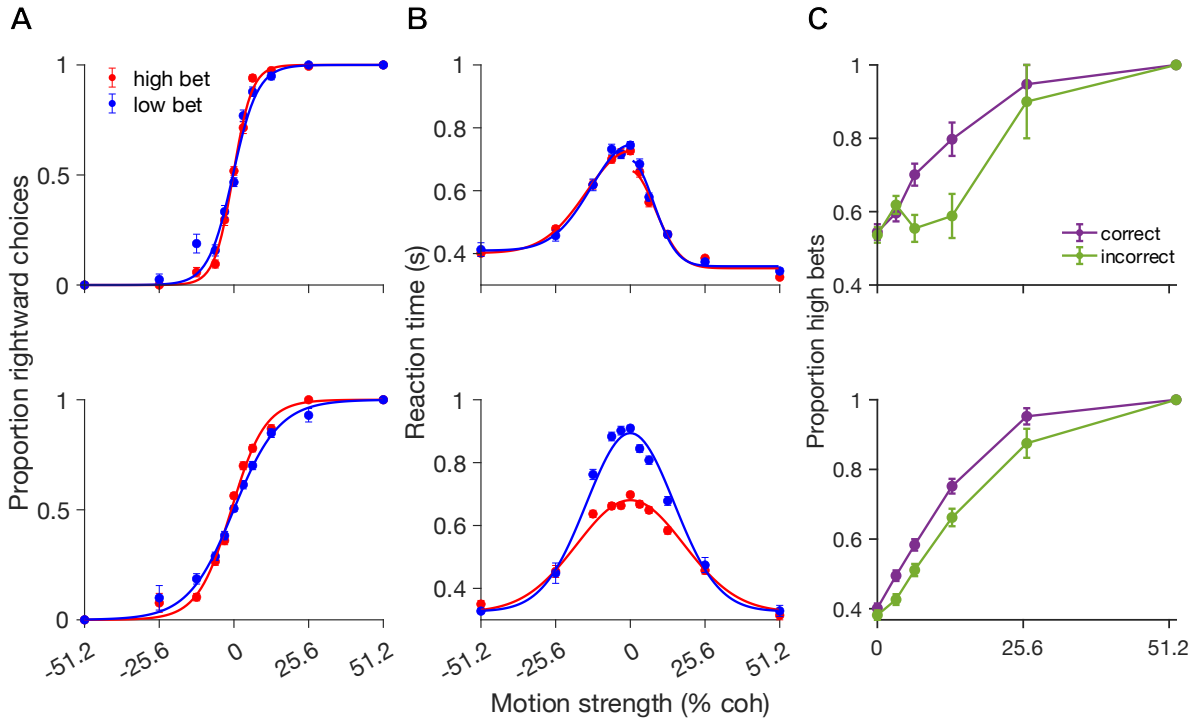

Figure S1. Three behavioral variables as a function of motion strength (% coh), including only the identical random seed trials ('double-pass' procedure (Bondy et al., 2018) adapted for 4 options), to control for motion energy within each coherence. (A) Proportion of rightward choices as a function of motion strength. Both monkey H (top) and monkey G (bottom) showed significant differences in sensitivity between high- and low-bet trials. Smooth curves are from logistic regression. (B) Mean RT as a function of motion strength. Monkey H (top) displayed no differences in RT between high and low wager when controlling for motion energy, whereas monkey G (bottom) still showed large differences for all but the highest two coherences. Smooth curves show Gaussian fits. (C) Proportion of high bets as a function of motion strength. Both monkeys bet high more frequently as motion strength increased, irrespective of accuracy. Error bars in all plots are +/- SE.

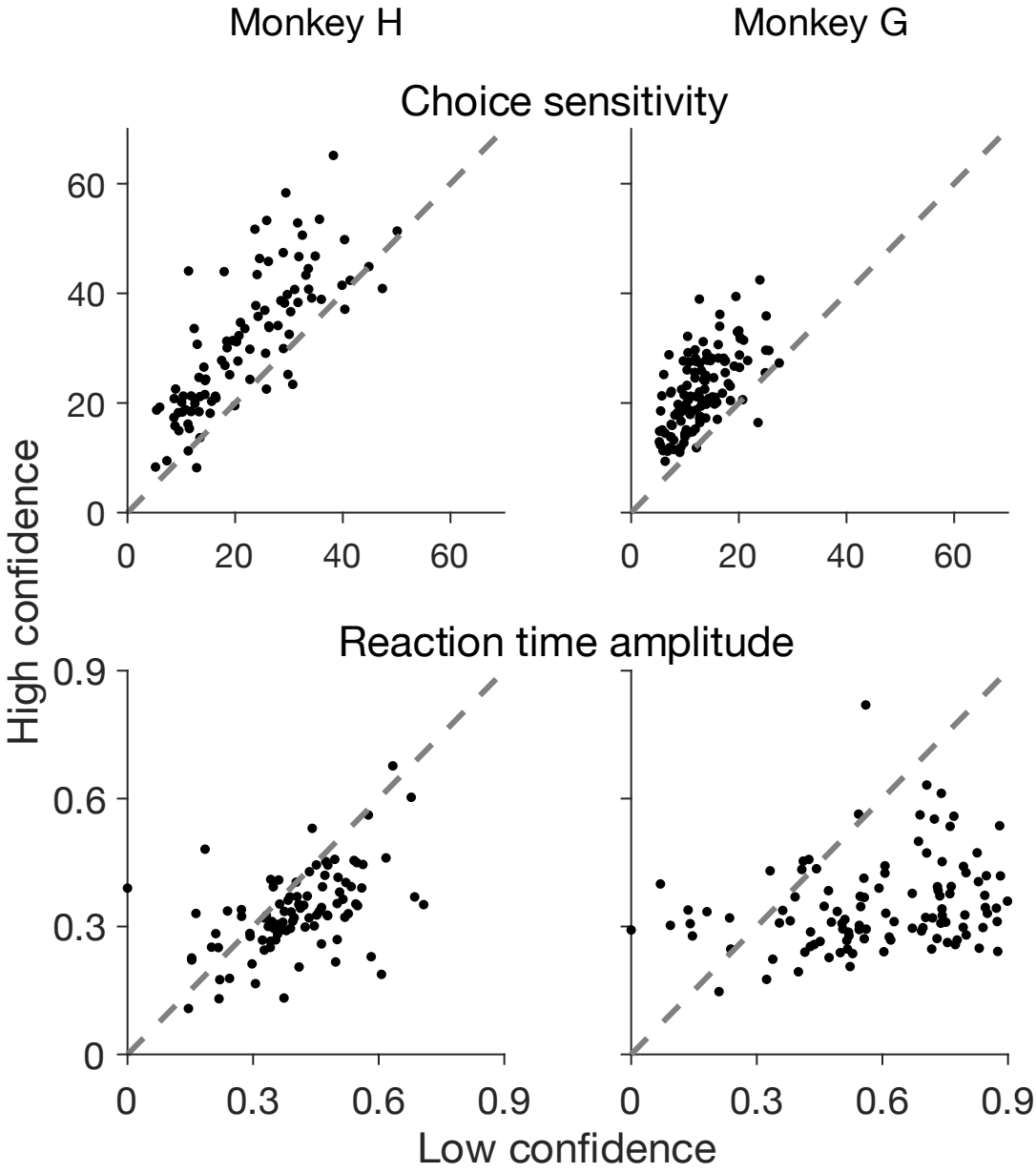

Figure S2. Session-by-session sensitivity and amplitude parameters from fitting logistic and Gaussian functions to choice and RT, respectively. The vast majority of individual sessions showed greater sensitivity (accuracy) and faster RT amplitude when the monkey bet high vs. low.

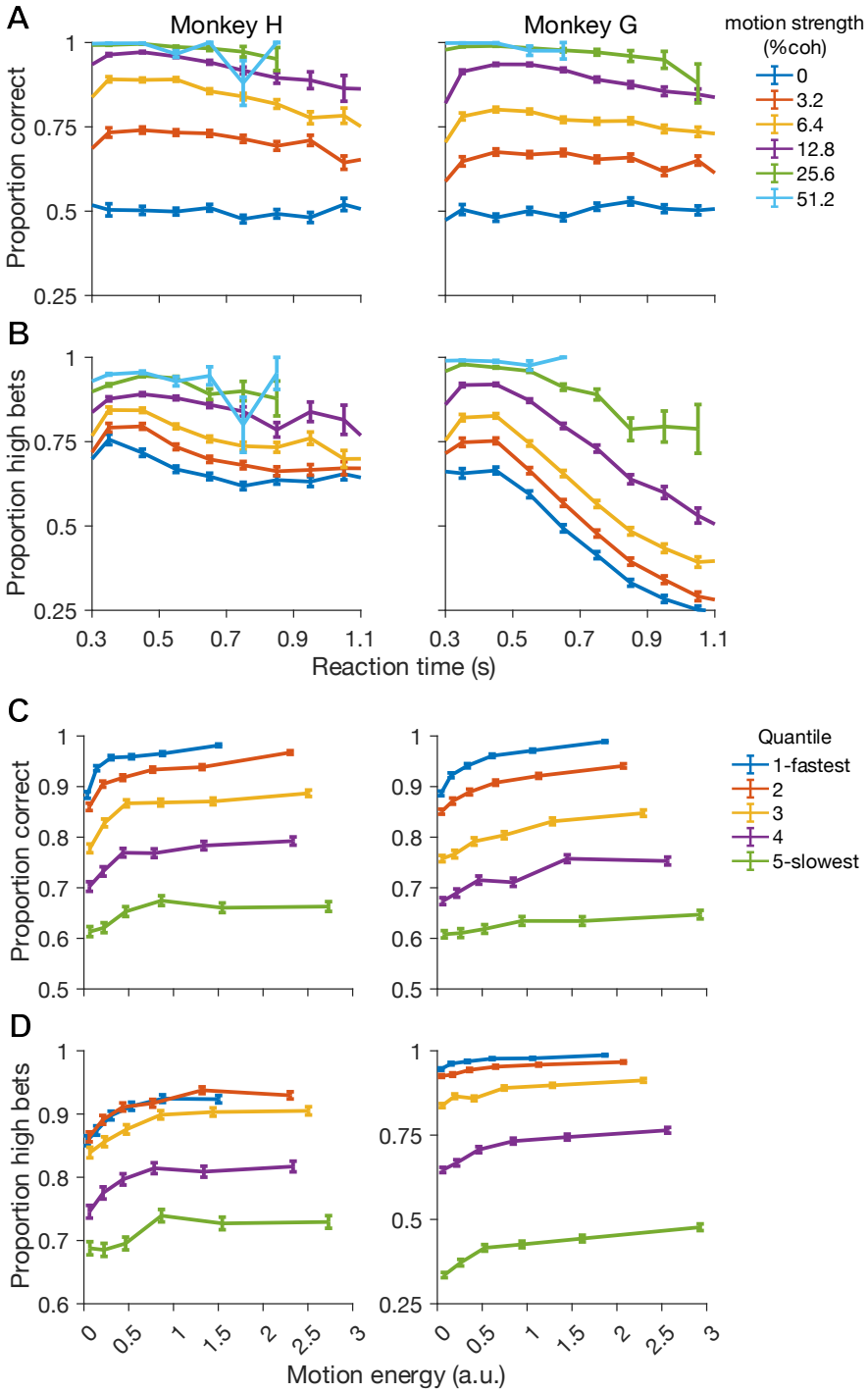

Figure S3. (A) Accuracy as a function of RT quantile, split by motion strength (% coh). Accuracy tended to decrease as a function of RT ($p < .0085$ for every coherence except 0% in both monkeys and 51.2% in monkey G). (B) Proportion of trials with a high bet as function of RT. Colors same as A. (C) Accuracy as a function of motion energy. Colors represents five different RT quantiles. Significant increases in accuracy were observed across all RT quantiles for both monkeys (p<0.01). (D) Proportion of trials with a high bet as a function of motion energy. Colors same as C.

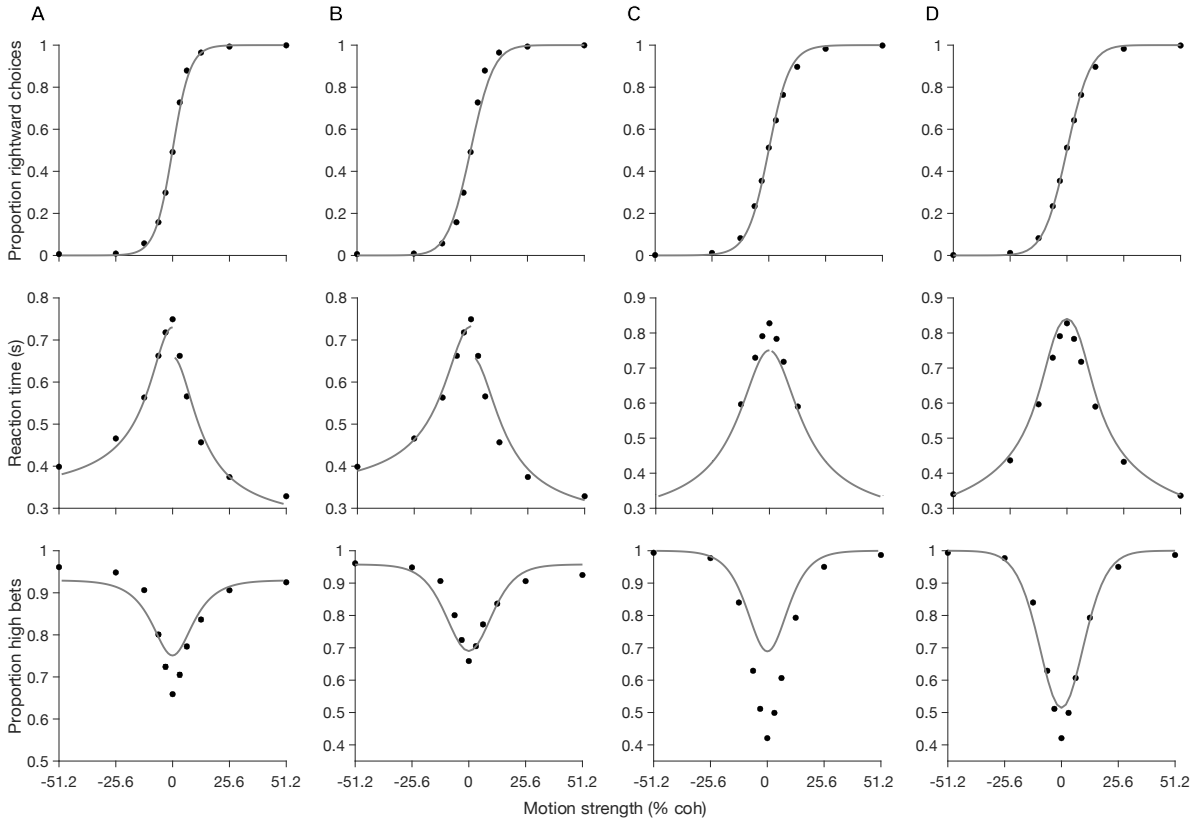

Figure S4. Serial and parallel model fits on unconditioned data. (A,C) Serial model fitted to choice, RT, and wager as a function of motion strength (% coh). (B,D) Parallel model fitted to choice, RT, and wager as a function of motion strength (% coh). First two columns correspond to monkey H and the last two to monkey G.

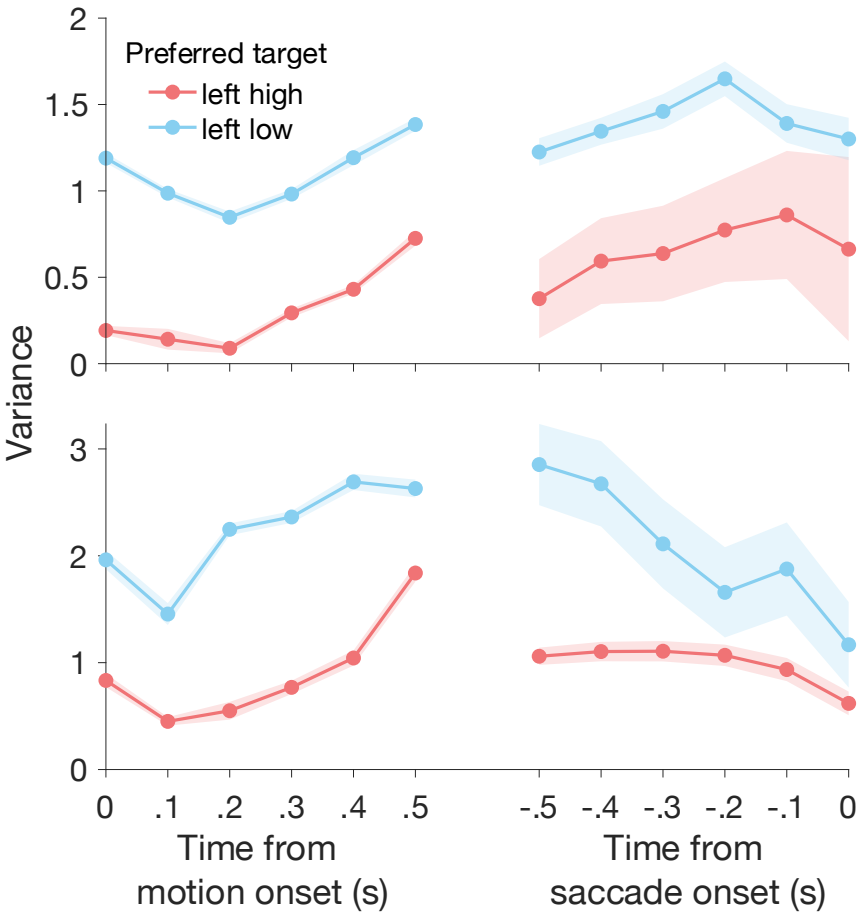

Figure S5. Variance of the conditional expectation (VarCE) estimated for the left-high and left-low populations of neurons. In both monkeys (H=top, G=bottom) VarCE begins to increase at approximately 0.2 s from motion onset, then decreases near saccade onset. Colors represent the two populations, and the shaded regions are standard errors calculated using a bootstrap.