

Neural Mechanisms of Post-error Adjustments of Decision Policy in Parietal Cortex

Highlights

- Humans and monkeys slow down after perceptual errors (post-error slowing)
- Post-error slowing arises from reduced sensitivity and increased decision bound
- Changes in sensitivity and bound are implemented in parietal response dynamics
- A balance of these dynamics stabilizes post-error accuracy at the expense of time

Authors

Braden A. Purcell, Roozbeh Kiani

Correspondence

roozbeh@nyu.edu

In Brief

Purcell and Kiani show that humans and monkeys slow after errors because of decreased sensitivity to sensory stimuli and increased decision bound. Electrophysiological recordings in parietal cortex reveal the neural mechanisms underlying these changes.

Neural Mechanisms of Post-error Adjustments of Decision Policy in Parietal Cortex

Braden A. Purcell¹ and Roozbeh Kiani^{1,*}

¹Center for Neural Science, New York University, New York, NY 10003, USA

*Correspondence: roozbeh@nyu.edu

<http://dx.doi.org/10.1016/j.neuron.2015.12.027>

SUMMARY

Humans often slow down after mistakes (post-error slowing [PES]), but the neural mechanism and adaptive role of PES remain controversial. We studied changes in the neural mechanisms of decision making after errors in humans and monkeys that performed a motion direction discrimination task. We found that PES is mediated by two factors: a reduction in sensitivity to sensory information and an increase in the decision bound. Both effects are implemented through dynamic changes in the decision-making process. Neuronal responses in the monkey lateral intraparietal area revealed that bound changes are implemented by decreasing an evidence-independent urgency signal. They also revealed a reduction in the rate of evidence accumulation, reflecting reduced sensitivity. These changes in the bound and sensitivity provide a quantitative account of choices and response times. We suggest that PES reflects an adaptive increase of decision bound in anticipation of maladaptive reductions in sensitivity to incoming evidence.

INTRODUCTION

The ability to change strategies when our actions fail to produce desired outcomes is fundamental to flexible behavior and survival in a complex, varying environment. Following errors, humans often slow down to respond on subsequent decisions, a phenomenon known as post-error slowing (PES) (Rabbitt and Rodgers, 1977). PES is a prime example of flexible changes in decision making and is widely assumed to play an adaptive role by preventing future errors (Botvinick et al., 2001; Dutilh et al., 2013; Goldfarb et al., 2012; Holroyd et al., 2005; Laming, 1979). However, the increased response times (RTs) in PES are rarely accompanied with tangible improvements in accuracy, suggesting that the phenomenon is more complex than a simple change in decision criterion (Danielmeier and Ullsperger, 2011; Gehring et al., 1993; Notebaert et al., 2009; Schroder and Moser, 2014). Understanding the mechanisms of PES and its role in adaptive decision making requires a systematic and quantitative study of the behavior and neural mechanisms involved in the decision-making process.

Previous investigations into the neural mechanisms of PES have focused almost exclusively on one aspect of this process: error detection and performance monitoring by medial frontal cortex (Schall et al., 2002). Error-related signals from medial frontal cortex correlate with slowing on the next trial (Gehring et al., 1993), and inactivation or lesions of medial frontal cortex can reduce or eliminate PES (di Pellegrino et al., 2007; Narayanan et al., 2013). However, far less is known about how the detection of errors in medial frontal regions affects future decision-making processes. Neuroimaging studies have revealed a diversity of changes in blood oxygen level-dependent (BOLD) responses across cortex following errors (Danielmeier and Ullsperger, 2011; King et al., 2010), but the limited spatial and temporal resolution of these techniques precludes model-based hypothesis testing for changes in the neural mechanisms of decision-making at the level of single neurons.

Intra-cortical recordings in non-human primates have excellent temporal and spatial resolution and, thus far, have revealed a network of posterior parietal, lateral prefrontal, and subcortical brain areas that work in concert to form decisions by integration of sensory information over time toward a criterion level (decision bound) (Hanes and Schall, 1996; Kiani et al., 2014b; Ratcliff et al., 2003; Roitman and Shadlen, 2002). Medial frontal cortex sends projections to several areas in this network (Huerta and Kaas, 1990) but does not directly represent the evidence accumulation process (Purcell et al., 2012; Stuphorn et al., 2010). Recent studies have demonstrated that integration of sensory evidence in this network is influenced by the behavioral state of the animal to adjust decision speed and accuracy (Hanks et al., 2014; Heitz and Schall, 2012), but it is unclear whether similar mechanisms extend to PES. Particularly, no previous study has quantified changes of neural responses following perceptual errors in these areas.

We measured PES in humans and monkeys trained to perform a motion direction discrimination task, establishing a non-human primate model for studying neural mechanisms of PES. By using quantitative models of the decision-making process, we hypothesized that PES in both species arises from a mixture of two effects: an increase in the decision bound and a decrease in the signal-to-noise ratio (SNR) of incoming sensory evidence. Together, these effects explain why PES may or may not be accompanied by changes in accuracy. We tested this hypothesis by recording from responses of single neurons in the lateral intraparietal area (LIP) of monkeys performing the task. Consistent with our model predictions, neural

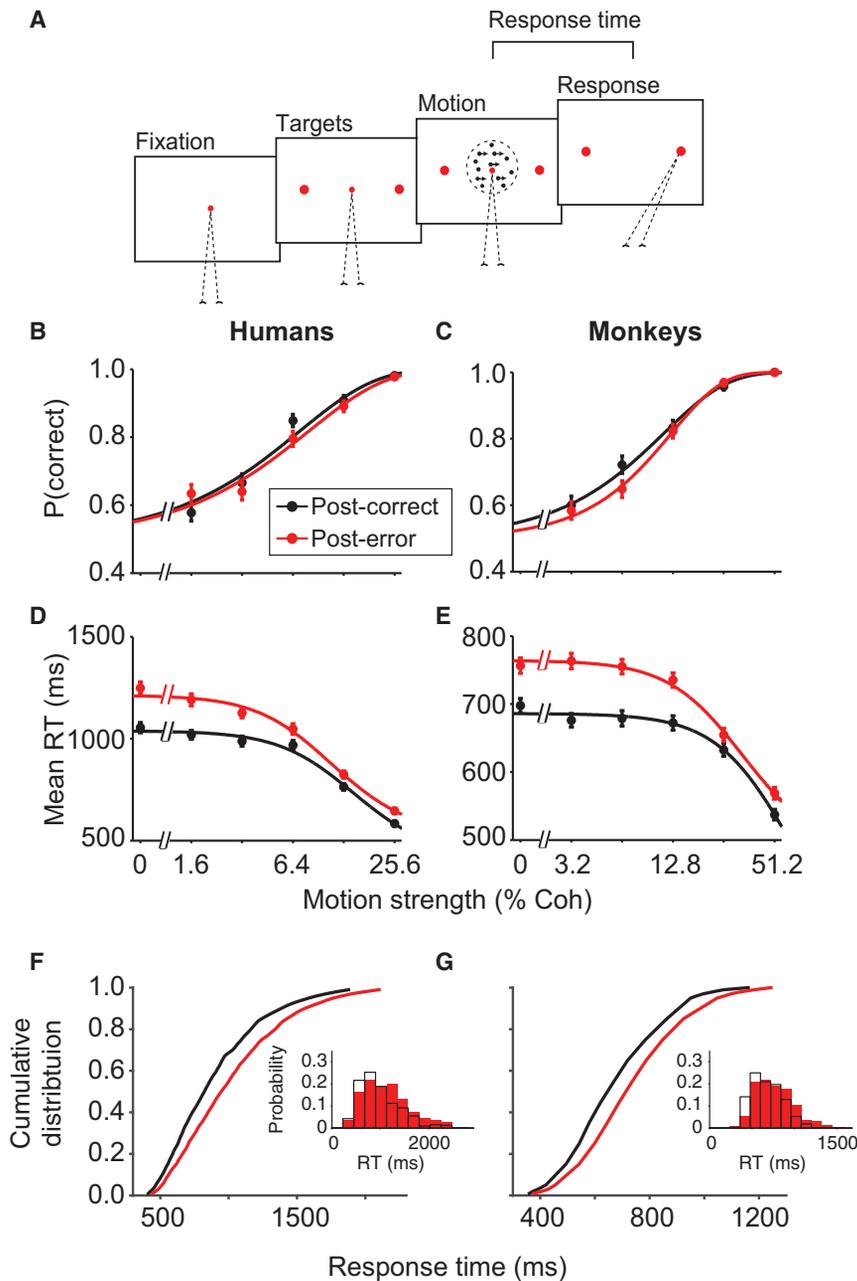


Figure 1. PES in Humans and Monkeys during Perceptual Decision Making

(A) Motion discrimination task. Subjects initiated a trial by fixating on a central point. The targets appeared after a short delay, followed by the random dot stimulus. When ready, subjects indicated the motion direction with a saccadic eye movement to a choice target. The motion strength (percentage of coherently moving dots [%Coh]) varied randomly across trials.

(B and C) Psychometric functions did not show appreciable change following negative feedback. Mean proportion correct as a function of motion strength for post-correct (black) and post-error (red) trials for six humans (B) and two monkeys (C) who exhibited PES. Smooth curves are fits to a logistic function. Error bars denote SEM. See Figures 8, S2, and S5 for behavioral data from subjects that did not exhibit PES.

(D and E) RTs are longer following errors. Smooth curves are fits to a hyperbolic tangent function.

(F and G) RT distributions show heavier upper tails following errors. Cumulative RT distributions for post-correct (black) and post-error (red) trials averaged across motion strengths demonstrate greater slowing for long RTs compared with short RTs. Insets show the RT probability density for an example motion strength (3.2% coherence) following correct (black open bars) and error (red solid bars) trials. Similar changes were evident for other motion strengths.

RESULTS

Humans and monkeys (*Macaca mulatta*) were trained to perform a reaction-time version of a motion direction discrimination task (Figure 1A). In each trial, the subject viewed a dynamic random-dots stimulus (Britten et al., 1992) and, when ready, reported the perceived motion direction with a saccadic eye movement. We manipulated task difficulty by varying the percentage of coherently moving dots (motion strength) on a trial-by-trial basis (see Experimental Procedures).

All subjects exhibited improved accuracy for stronger motion (Table S1; Figures 1B and 1C; Equation 1: humans, $\beta_1 = 20.85 \pm 1.01$, $p < 10^{-10}$; monkeys, $\beta_1 = 11.91 \pm 0.46$, $p < 10^{-10}$), compatible with previous studies (Roitman and Shadlen, 2002). Also as expected, RTs decreased sharply for stronger motion (Table S1; Figures 1D and 1E; Equation 2: humans, $\beta_1 = -2,005 \pm 104.56$, $p < 10^{-10}$; monkeys, $\beta_1 = -267 \pm 15.00$, $p < 10^{-10}$). All subjects were trained to a high level of performance as evidenced by their low psychophysical thresholds (ranges of 8.2%–13.0% for monkeys and 4.6%–9.9% for humans). On average, monkeys were faster than humans (mean RTs of 604 ± 0.9 ms for monkeys and 915 ± 3.6 ms for humans), but humans had lower thresholds than monkeys. Therefore, we used a slightly more difficult

responses in LIP revealed both decreased sensitivity and an increase in decision bound after errors. Neural responses also showed that both factors were implemented through changes in the dynamics of decision formation. Sensory and motor delays, sensory and accumulation noise, and static components of neural responses (start point and endpoint) remained unchanged, ruling out competing hypotheses that could not be distinguished solely on the basis of the behavior and computational models. These results indicate that changes of neural response dynamics after errors implement rapid adjustments of the decision policy by elongation of decision time in order to mitigate reduction of accuracy.

responses in LIP revealed both decreased sensitivity and an increase in decision bound after errors. Neural responses also showed that both factors were implemented through changes in the dynamics of decision formation. Sensory and motor delays, sensory and accumulation noise, and static components of neural responses (start point and endpoint) remained unchanged, ruling out competing hypotheses that could not be distinguished solely on the basis of the behavior and computational models. These results indicate that changes of neural response dynamics after errors implement rapid adjustments of the decision policy by elongation of decision time in order to mitigate reduction of accuracy.

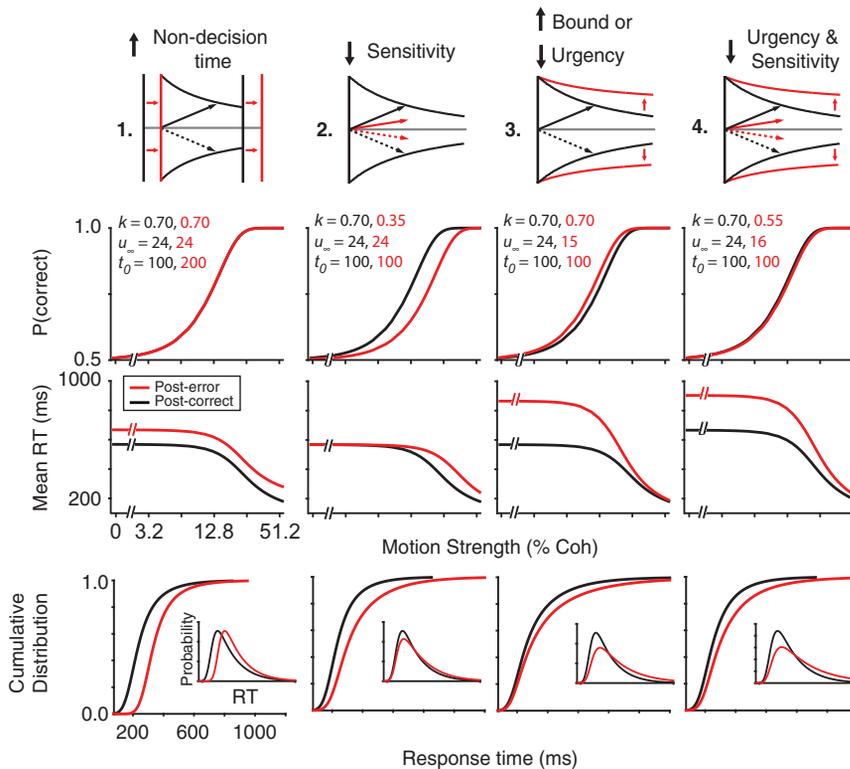


Figure 2. Hypothesized Mechanisms of PES and Their Behavioral Predictions

We used a DDM to predict expected changes in post-error behavior under different hypotheses for the mechanism of PES. Each column shows one of these hypotheses, which can be distinguished on the basis of changes in the psychometric and chronometric functions and the shape of the RT distributions. The first row shows a schematic of these hypotheses. The second and third rows illustrate predictions for psychometric and chronometric functions, respectively. The final row shows predicted changes in cumulative RT distributions (insets show the probability density function for an example motion strength). The sensitivity (k), urgency (u_{∞} ; Equation 4), and non-decision time (t_0) parameters used for the model simulations are shown in row two. All other parameters were constant for post-correct and post-error conditions.

stimulus set for humans to approximately match the error rates across species (mean percentages correct of 77.2 ± 0.20 for monkeys and 74.8 ± 0.33 for humans).

Because the motion strength and direction were independent across trials, the optimal strategy was to make the decision solely on the basis of the sensory information provided on each trial. Nonetheless, both humans and monkeys were influenced by the outcome of previous trials. The majority of subjects (73%; six humans and monkeys O and N) exhibited longer RTs in trials following error feedback (Table S1; mean human $\Delta RT = 117.0 \pm 14.82$ ms; monkeys O and N mean $\Delta RT = 59.5 \pm 6.75$ ms), compatible with past studies of PES. To maintain our focus, we first report results from this majority group that exhibited clear PES. Later in this paper, we report results from the minority that responded differently to errors.

Quantitative Hypotheses of PES

Electrophysiological recordings of neural responses from parietal cortex, frontal cortex, basal ganglia, and superior colliculus of monkeys engaged in perceptual decisions have shown that integration of noisy sensory evidence over time toward a decision bound underlies the choice and RT (Hanes and Schall, 1996; Kiani et al., 2014b; Ratcliff et al., 2003; Roitman and Shadlen, 2002). Many of the recorded neurons represent the integral of evidence in favor of the target in their response fields (RFs). Furthermore, their firing rates reach a common threshold shortly before a saccade toward RF. A simple process model that closely approximates these responses consists of accumulators that integrate noisy evidence toward a decision bound (Gold and Shadlen, 2007). The accumulator that reaches its bound first dic-

ates the choice and the time to bound indicates the decision time. Variants of this “bounded accumulation” model have been shown to provide quantitative explanations of behavior across a large array of decisions (Bogacz et al., 2006; Busemeyer and Townsend, 1993; Link, 1992; Usher and McClelland, 2001). To study the mechanisms of PES quantitatively, we used a simple variant of this model, the drift-diffusion model (DDM), in which the two competing accumulators are replaced with a single accumulator with an upper and a lower bound that correspond to the two options (Gold and Shadlen, 2007; Link, 1992). We adopted the DDM because of its simplicity, but other variations of the bounded-accumulator framework will produce identical conclusions.

Our model offers a tool to formalize longstanding hypotheses about the neural mechanisms of PES (Figure 2). A first hypothesis is that PES occurs because subjects postpone the onset of evidence accumulation in order to recover from the error (Laming, 1979; Pouget et al., 2011; Rabbitt and Rodgers, 1977). In the DDM, this would be captured by the non-decision time parameter that reflects the sum of sensory and motor delays. If PES arises from error-induced elongation of non-decision time, slowing would be equivalent at all motion strengths, and the proportion correct would remain unchanged (Figure 2, first column). Furthermore, the RT distributions should shift in time but maintain their shape for each motion strength.

A second hypothesis is that PES arises from impairment of the subject’s sensitivity to incoming evidence following errors, perhaps because the negative feedback or unexpected outcome distracts and diverts attention (Notebaert et al., 2009) by influencing either sensory cortex (Treue and Martinez Trujillo, 1999) or transmission of information from sensory cortex to decision-making areas (Green et al., 2010). This hypothesis is plausible because accuracy rarely increases and occasionally decreases following errors (Danielmeier and Ullsperger, 2011). In the DDM, this is captured by a sensitivity parameter that defines

how the evidence that the decision maker extracts from the stimulus scales with stimulus strength (Gold and Shadlen, 2007). Lower sensitivity reduces available evidence per unit time and slows down evidence accumulation, causing longer RTs (Figure 2, second column). The increased RTs, however, are also associated with reduced accuracy because of the reduced reliability of evidence (Figure 2, second column). Moreover, slowing of RTs should be maximal for higher motion strengths because the rate of evidence accumulation (sensitivity \times motion strength) is changed more for those stimuli. Finally, the RT distribution should be shifted and only modestly skewed for each motion strength.

A third hypothesis is that PES is caused by an adaptive increase in the decision bound, which increases the total evidence that must be accumulated before a choice is made. This hypothesis, which is favored by many leading models of executive control and adaptive behavior (Holroyd et al., 2005; Laming, 1979), predicts increased accuracy at the cost of longer RTs (Hanks et al., 2014; Heitz and Schall, 2012). Changes in bound height also predict larger effects on RT when the rate of evidence accumulation is low, causing maximal slowing at low motion strengths (Figure 2, third column) and skewing the RT distribution of each motion strength by creating a heavier upper tail.

A change in decision bound could be implemented through different neural mechanisms. Two obvious possibilities include changes in the final firing rate or the initial firing rate of the neurons that represent evidence accumulation. A less obvious but computationally effective alternative is to adjust the total evidence needed to reach the bound through a dynamic stimulus-independent rise of the accumulators, referred to as urgency (Churchland et al., 2008; Standage et al., 2011; Thura et al., 2014). Because urgency equally increases the firing rates of neurons representing all potential choices, it reduces the total evidence that must be accumulated for committing to a choice, as if the bound had changed although the start points and endpoints remain fixed. In the DDM, changes in the start point and endpoint are represented by static shifts in bound heights, but urgency is implemented with a dynamic collapse of the bounds (Hanks et al., 2014). Note that the DDM bound collapse in this paper is solely a mathematical shortcut to capture urgency in a simplified model that accumulates evidence along a single dimension. Also note that both static bound changes and changes in urgency make qualitatively similar predictions about accuracy and RTs in PES (Figure 2, third column). In our initial modeling, we use a DDM with urgency, and later we will distinguish changes in urgency from static changes in start point and endpoint through quantitative model comparisons and, more important, analyses of neural responses.

The aforementioned mechanisms are not mutually exclusive. In fact, it is likely that PES results from changes in multiple components of the decision-making process. For example, sensitivity may be reduced by unexpected negative feedback, and simultaneously the decision bound is raised to compensate. The model can disentangle these mixed effects into a combination of reduced urgency and sensitivity (Figure 2, fourth column). The opposite effects of the two changes on accuracy could balance out, leading to little or no change in psychometric functions. However, the RT distributions develop a heavier tail, and the

mean RTs increase for all stimuli, especially in trials with lower motion strengths.

PES Is Due to Decreased Urgency and Sensitivity

To distinguish different PES mechanisms, we characterized the post-correct and post-error performance of our subjects. Human and monkey behavior after errors exhibited common trends, suggesting that PES in both species reflects a common set of neural mechanisms (Figure 1). The increase in mean RT following errors was significantly larger for lower motion strengths, for both humans (Figures 1D and S1; Table S1; Equation 2; $\beta_3 = -512.9 \pm 144.06$, $p = 3.7 \times 10^{-4}$) and monkeys (Figure 1E; Equation 2; $\beta_3 = -94.4 \pm 26.31$, $p = 3.4 \times 10^{-4}$). Also, for each motion strength, PES was maximal for longer RTs, causing a heavier upper tail in the RT distribution (humans: Figure 1F; Equation 3; $\beta_2 = 21.95 \pm 2.92$, $p < 10^{-10}$; monkeys: Figure 1G; Equation 3; $\beta_2 = 5.88 \pm 1.31$, $p = 1.6 \times 10^{-5}$). Finally, despite the RT changes, accuracy did not show an appreciable difference following errors, for either humans (Figures 1B and S1; Table S1; Equation 1; $\beta_2 = -2.31 \pm 1.37$, $p = 0.09$) or monkeys (Figure 1C; Equation 1; $\beta_2 = -0.63 \pm 0.99$, $p = 0.53$). The pattern of RT changes together with the constancy of accuracy rules out isolated changes in sensory-motor delays, sensitivity, and bound height as the cause of PES. Rather, PES appears to be caused by a mixture of changes in the decision-making processes.

To pinpoint the mechanisms underlying PES, we fit a DDM to the RT distributions of individual subjects following correct and error feedback (see Experimental Procedures). Figures 3 and S1 demonstrate that the model reproduces all key aspects of PES. The fits successfully accounted for the full RT distributions across all motion strengths (Figure S1, third and fourth columns; R^2 values for the match between the model and data for individual subjects ranged from 0.86–0.98). They also captured the properties of PES, showing increased slowing at lower motion strengths (Figures 3C, 3D, and S1, first column) and also increased skewness (heavier tail) in RT distributions of each motion strength (Figures 3E, 3F, and S1, comparison of third and fourth columns).

The model can be independently verified by its ability to predict choices. Because we optimized the model parameters solely on the basis of RTs and irrespective of choice, we could generate predictions about the probability of each choice. Those predictions accurately matched individual subjects' choice probabilities (R^2 range 0.81–0.98 for psychometric functions), in line with the experimental observation that accuracy is unchanged following errors, although this was not forced by the model (Figures 3A, 3B, and S1, second column). This successful prediction supports that the model is a reliable approximation of the neural mechanisms that underlie choice and RT.

The best fitting model parameters offer strong predictions about the neural mechanisms of PES. Figure 4 illustrates those predictions by comparing the model parameters for post-error and post-correct trials (see Table S2 for all model parameters). First, urgency was reduced following errors for all humans (Figure 4A) and monkeys (Figure 4B) exhibiting PES. This reduction resulted in a significant elevation of the mean bound height (Figure 4C; Equation S2; $\beta = 1.91 \pm 0.402$, $p = 2.1 \times 10^{-3}$). Second, the stimulus sensitivity decreased for both monkeys and humans

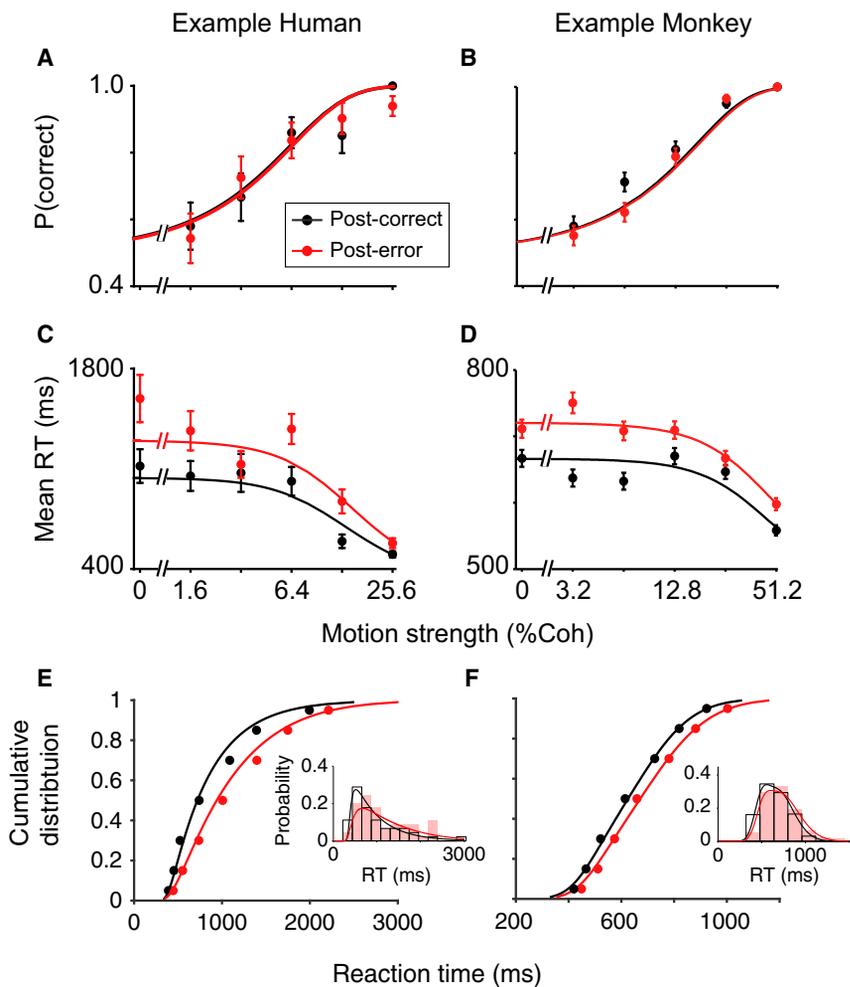


Figure 3. The DDM Fit to RT Distributions Explains PES and Predicts No Change in Accuracy

Model fits and predictions for one representative human subject (left) and one representative monkey subject (right). See [Figures S1](#) and [S2](#) for a detailed summary of model results for all subjects. (A and B) Observed (solid points) and predicted (lines) proportion correct as a function of motion strength for a representative human (S6; first column) and monkey (O; second column). (C and D) Observed and expected mean RT as a function of motion strength.

(E and F) Observed (solid points) and fitted (lines) cumulative RT distributions averaged across motion strengths. Insets show the corresponding observed (bars) and fitted (lines) RT probability density functions for an example motion strength (3.2% coherence). For both species, the model fits the heavier upper tail (greater slowing at larger RTs compared to short RTs), ruling out pure non-decision time changes ([Figure 2](#), column 1). Similar results were obtained for all motion strengths. Error bars denote SEM.

represent, respectively, the start point of evidence accumulation and the span over which the firing rates change during decision formation.

LIP responses did not show any change in the static features of the ramp: onset time, start point firing rate, and endpoint firing rate. The mean onset time of the ramp was indistinguishable for post-error (120 ± 21.3 ms) and post-correct (119 ± 17.6 ms) trials ($p = 0.66$, Wilcoxon

signed-rank test; [Figure 5B](#)), ruling out increased sensory delay as a main cause for PES. LIP neurons also showed similar firing rates at the beginning of the integration process following errors ([Figure 5C](#); $F = 0.01$, $p = 0.91$, mixed-measures ANOVA) and achieved the same common level before the saccade ([Figure 5D](#); $F = 0.15$, $p = 0.70$). Finally, we verified that the total excursion (endpoint to start point) was indistinguishable for post-correct and post-error trials ([Figure 5E](#); $F = 0.15$, $p = 0.69$). Identical trends were evident in the normalized responses, which reduced across-neuron variation in total excursion (see [Experimental Procedures](#)). Overall, the onset time and the range of the ramp remained stable.

Dynamics of LIP Neural Responses Change with PES

[Figures 5A](#) and [S3](#) show the mean firing rates of 114 neurons from monkeys O and N. After the stimulus onset, there was a gradual ramping of firing rates. Responses increased when the monkey ultimately chose the target in the RF (T_{in}) and declined when the choice was in the opposite direction (T_{opp}). The “buildup rate” depended on the motion strength: stronger motion toward T_{in} increased buildup rates ([Figure S3](#); Equation 5; $\beta_2 = 1.6 \pm 0.13$, $p < 10^{-10}$), and stronger motion toward T_{opp} decreased buildup rates (Equation 5; $\beta_3 = -0.4 \pm 0.13$, $p = 8.6 \times 10^{-4}$), consistent with the accumulation of different levels of sensory evidence. Also, as expected by the model, the firing rates reflected the decision commitment by converging to a common level at the end of the ramp ~ 70 ms prior to the T_{in} saccade ([Figure S3](#)). The start point and endpoint of the ramp

However, PES was associated with a change in the dynamics of neural responses between the start point and endpoint. This change consisted of two components: a stimulus-independent component that corresponded to a reduced urgency and a stimulus-dependent component that corresponded to a reduced sensitivity to stimulus strength. Below we expand on these components.

Following [Churchland et al. \(2008\)](#), we estimated the form of the urgency signal by averaging firing rates across T_{in} and T_{opp} trials. Because we combined trials in which evidence is for and against a particular response, the effect of incoming evidence

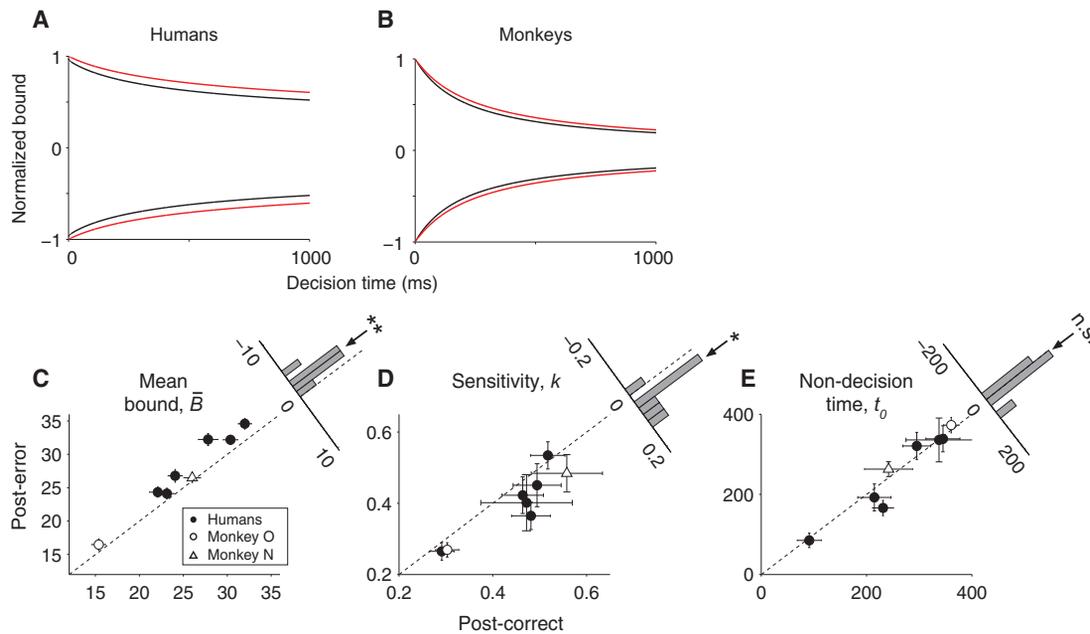


Figure 4. Joint Reductions in Urgency and Sensitivity Explain PES

(A and B) Decision bound for post-correct (black) and post-error (red) trials averaged across humans (A) and monkeys (B) that exhibited PES. To ensure all subjects contributed equally to the averages, we normalized the bound heights by dividing by each subject's maximum across all conditions and times. Urgency is reduced after errors for both species, resulting in an overall greater bound height.

(C–E) Mean bound height (C), sensitivity (D), and non-decision time (E) for post-error trials versus post-correct trials. Mean bound height was computed for each subject as the average bound over the middle 95% of the decision time distribution. Error bars denote SE. Histograms show the distribution of the difference of post-error and post-correct model parameters across subjects. Arrows over the histograms indicate the mean difference (Equation S2, ** $p < 0.01$, * $p < 0.05$; n.s., not significant).

cancels out, and only the nonspecific urgency modulation remains. Figure 6A shows these neurally derived estimates, confirming the presence of a rising urgency signal that was well characterized by the same hyperbolic function used in the DDM. It also shows that the urgency was reduced following errors. Across individual neurons, there was a significant reduction in urgency following errors, as captured by a decrease in the asymptote of the best-fitting hyperbolic function (Equation S3; $\Delta u_{\infty} = -0.60 \pm 0.20$, $p = 1.7 \times 10^{-2}$, Wilcoxon signed-rank test), while the rise time and start point did not change ($\Delta \tau_{1/2} = -6.9 \pm 50.41$, $p = 0.84$; $\Delta \beta_0 = 0.016 \pm 0.012$, $p = 0.99$). A similar difference in post-error and post-correct urgency was observed when the dynamic component of the urgency signal was estimated using an alternative method developed by Hanks et al. (2011) (Figure 6B; see Supplemental Experimental Procedures). Put together, the unchanged start point and endpoint firing rates and the reduced urgency in LIP responses implemented a dynamic increase in the decision bound and augmented the overall evidence that subjects accumulated for decisions. We found further support for this conclusion through nested model testing. Replacing the dynamic urgency (bound collapse) in the model with a static change in bound height significantly reduced the quality of the fits in the majority of subjects ($p < 0.05$, likelihood ratio test in seven of eight subjects exhibiting PES).

LIP neural responses also reflected the reduction of stimulus sensitivity. Recall that our model suggested reduced sensitivity as one of the two factors that induced PES. A combination of ur-

gency and sensitivity reductions makes specific predictions for changes in the buildup rate of LIP neurons (Figure 7A): an overall reduction of buildup rates for T_{in} motion directions, smaller or no effects for T_{opp} motion directions, and shallower dependency of buildup rates on motion strength. LIP responses were consistent with this prediction (Figure 7B). The lack of change in the T_{opp} buildup rates could potentially be due to a floor effect (firing rates cannot go below zero). Importantly, however, we found that the slope of buildup rate as a function of motion strength for T_{in} motion directions was significantly shallower following errors (Figure 7B; Equation 5; $\beta_4 = -0.46 \pm 0.18$, $p = 9.0 \times 10^{-3}$). This effect was still clear even when we regressed out the relationship of RT and buildup rate (Equation 5; $\beta_4 = -0.35 \pm 0.17$, $p = 3.9 \times 10^{-2}$), indicating that RT differences alone cannot explain the result. Altogether, LIP responses are most consistent with a model in which both evidence-dependent (sensitivity) and evidence-independent (urgency) inputs to the accumulation mechanism are reduced following errors.

Decreased SNR with PES Is Not Explained by Elevated Noise

From a theoretical perspective, the combined reduction of urgency and sensitivity could be replaced with a combination of reduced urgency and increased noise without affecting the quality of RT fits. Note that the critical factor in the model is the SNR. Traditionally, noise is assumed to be constant, and drift rate and decision bound are expressed in units of noise (Link, 1992).

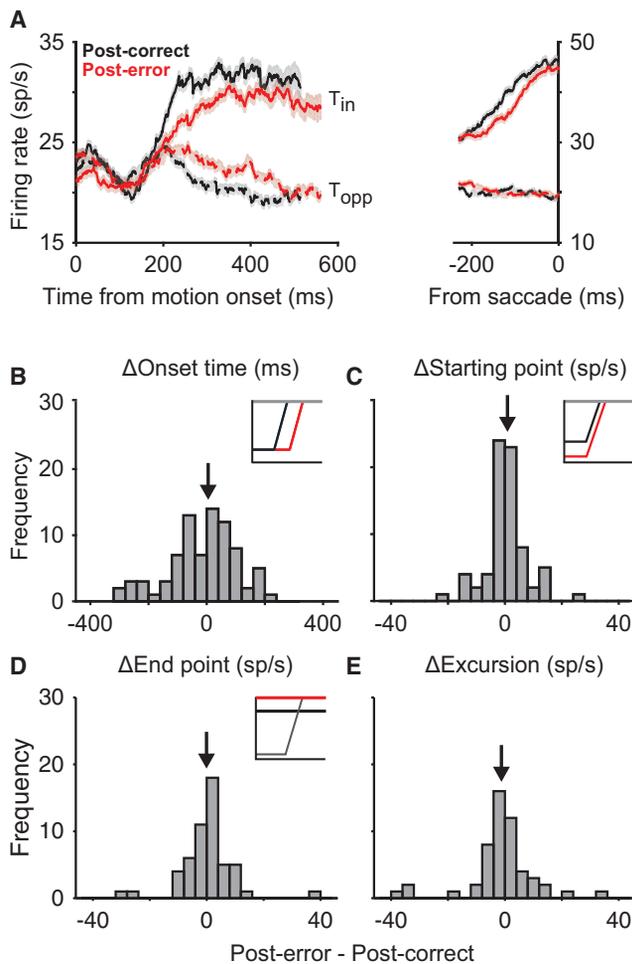


Figure 5. The Static Components of the Decision-Making Mechanism Do Not Change with PES

(A) Average responses of LIP neurons for monkeys O and N in post-correct (black) and post-error (red) trials. Solid and dashed lines show trials in which the choice was made to the target inside (T_{in}) or opposite (T_{opp}) the neuron's RF.

(B–E) There was no change in the static components of the LIP responses. The histograms show the difference between post-correct and post-error trials for the onset time of ramping responses (B); magnitude of neural responses at the beginning of evidence accumulation (C); magnitude of neural responses at the end of accumulation, prior to the saccade (D); and total excursion between the start point and endpoint (E). Insets illustrate schematic changes in activity that would be expected if the static components of the decision-making process explained PES. Black arrows indicate the mean of the distribution.

Assuming fixed noise, our model predicted a reduction of sensitivity. A more general conclusion from the model would be a reduced SNR, which could be implemented by reduced sensitivity, elevated noise, or both. To formally establish this point, we fit an alternative model in which we fixed sensitivity and allowed noise to vary across post-correct and post-error trials (equal number of free model parameters). For all eight subjects exhibiting PES, the quality of the fit was statistically unchanged (bootstrap test of model log likelihoods, $p > 0.46$ for all subjects). An examination of the best fitting parameters revealed that PES

could be equally well explained with a mixture of reduced urgency and elevated noise (Figure S4; Equation S2: \bar{B} : $\beta = 3.96 \pm 1.239$, $p = 1.5 \times 10^{-2}$; σ : $\beta = 0.12 \pm 0.038$, $p = 1.6 \times 10^{-2}$; t_0 : $\beta = -0.12 \pm 7.923$ ms, $p = 0.98$). Thus, these two models (noise change and sensitivity change) are indistinguishable solely on the basis of behavioral data, but we can turn to physiology to resolve this apparent model mimicry.

The LIP responses confirmed that reduced SNR is due at least partly to reduced sensitivity (above). They also enabled us to directly test whether PES could be partially attributed to elevated noise. Elevated noise manifests itself as increased trial-to-trial fluctuations of the rate function that underlies spiking activity. We quantified this rate fluctuation by computing the variance of conditional expectation (VarCE), a statistic that parses true across-trial variability in response rate from point-process variability (Equation S4; Churchland et al., 2011). For both post-correct and post-error trials, VarCE gradually increased over time following motion onset, consistent with the accumulation of noisy evidence, and reached nadir around the time of saccade, consistent with convergence to a bound. However, there was no significant increase in VarCE for post-error relative to post-correct trials (Figure 7C; bootstrap, $p > 0.05$ for all comparisons 100–400 ms after motion onset). Results were identical when we analyzed the Fano factor (spike count variance/spike count mean; Equation S5). Altogether, our data demonstrate that decreased SNR with PES is due primarily to reductions in sensitivity rather than elevated noise.

Post-error Speeding Is Associated with Reduced Sensitivity and Reduced Decision Bound

In the previous sections, we focused on the majority of subjects who exhibited clear PES. Here, we explore the remaining minority. Two subjects (one human, S7, and one monkey, D) consistently decreased their RTs following errors (Figures 8 and S2; Table S1). “Post-error speeding” has been previously reported, especially in fast-response regimes when subjects emphasize speed over accuracy (King et al., 2010; Notebaert et al., 2009). Our two subjects with post-error speeding presented an opportunity to contrast the mechanism of post-error speeding and PES. For both subjects, the slope of the psychometric function was significantly reduced after errors (Figures 8A and S2; Equation 1; $\beta_2 = -2.02 \pm 0.82$, $p = 0.01$), and post-error speeding was maximal at lower coherences (Equation 2; $\beta_3 = 161 \pm 52.77$, $p = 2.3 \times 10^{-3}$). This pattern of behavior was best explained with a combination of decreased sensitivity and, importantly, decreased decision bound (Table S2), perhaps to rush through trials with impaired sensitivity. Our nested behavioral models indicated the necessity of bound changes to explain post-error speeding for both subjects (monkey D, $p = 8.9 \times 10^{-8}$; subject S7, $p < 10^{-10}$) but could not distinguish between static bound changes or dynamic changes in urgency.

The responses of 37 LIP neurons in monkey D indicated a strong increase in the start point of integration (Figure 8C; $F = 8.89$, $p = 3.1 \times 10^{-3}$, mixed-measures ANOVA) and no change in the endpoint ($F = 0.04$, $p = 0.83$), leading to an overall reduction of the neural response excursion for this monkey ($F = 4.36$, $p = 3.7 \times 10^{-2}$). This large increase in start point obscured our ability to reliably compare neural urgency signals across conditions.

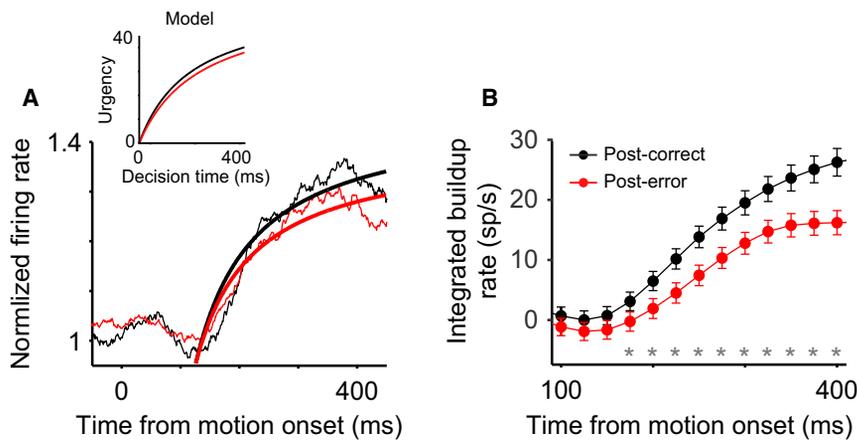


Figure 6. LIP Responses Reflect a Reduction of Urgency following Errors

(A) LIP responses averaged across all choices (T_{in} and T_{opp}) and stimuli from monkeys O and N in post-correct (black) and post-error (red) trials. Averaging across choices cancels out the effect of incoming sensory evidence such that only an evidence-independent urgency signal remains (Churchland et al., 2008). Smooth curves are fits of a hyperbolic function (Equation S3). Inset shows the average urgency signal based on the best fitting DDM parameters for the two monkeys (see Figure 4B).

(B) Neurally derived estimate of urgency based on the mean buildup rates computed during decision formation (Hanks et al., 2011). The mean buildup rate was computed in a running window (width = 150 ms, 25 ms steps) 100–400 ms after motion onset to obtain a piecewise estimate of the first

derivative of the urgency function. The integral of the piecewise buildup rates provides an estimate of the dynamic urgency signal. The urgency signal was estimated independently for post-correct (black) and post-error (red) trials. Asterisks indicate time bins in which the urgency signals on post-correct and post-error trials were significantly different ($p < 0.05$; parametric bootstrap, 1,000 samples). Error bars denote SEM.

Therefore, we cannot rule out changes in urgency in addition to changes in start point. Regardless of its origin, it seems clear that a reduced decision bound is highly likely to play a central role in shaping post-error speeding. A stronger conclusion about the underlying neural mechanism of this bound reduction must await data from additional monkeys.

Our last subject, monkey B, did not show noticeable changes of behavior following errors (Figures S5A and S5B; Table S1), offering a control to test whether the aforementioned changes in neural responses following errors were associated with behavioral changes or were merely incidental to negative feedback. Post-correct and post-error responses of 52 neurons recorded from this monkey were largely identical (Figure S5C). The buildup rates (Equation 5; $\beta_4 = 0.34 \pm 0.48$, $p = 0.48$), urgency ($\Delta u_\infty = 0.27 \pm 0.43$, $p = 0.63$; $\Delta\tau_{1/2} = -42.87 \pm 60.07$, $p = 0.35$, Wilcoxon signed-rank test), and overall excursion ($F = 0.002$, $p = 0.96$, mixed-measures ANOVA) were unchanged following errors. Altogether, these results indicate that changes in neuronal responses are specific to the type of post-error behavioral adjustment.

DISCUSSION

Seminal work by Laming (1979) has formed the basis for numerous computational frameworks that explain PES through adjustments in the decision bound (Goldfarb et al., 2012; Holroyd et al., 2005) and has led various clinical and neurophysiological studies to interpret PES as a direct index of adaptive behavior (e.g., Sergeant and van der Meere, 1988). However, a pure bound change makes a strong prediction that accuracy increases following errors (Figure 2), a prediction that has been contradicted by many studies, including this one (e.g., King et al., 2010). This has led some researchers to question whether PES reflects any adaptive process whatsoever or whether this interpretation should be exclusively reserved for the rare cases in which post-error accuracy is also elevated (Schroder and Moser, 2014).

Through a detailed analysis of behavior, we demonstrated that increases in decision bound are necessary to explain PES, even in the absence of elevated accuracy. The key experimental observations are more slowing at low stimulus strengths and heavier tails in the RT distribution. This pattern is best explained by a reduction in urgency that effectively raises the decision bound. Although our results support adaptive theories of PES, we note that this explanation is necessarily incomplete; some additional mechanism is needed to explain why accuracy rarely increases with PES.

The additional mechanism is a concurrent reduction in sensory SNR, which eliminates expected gains in accuracy with increased decision bound. Decreased SNR is maladaptive for the goals of the task: it reduces the reward rate by making responses both slower and less accurate and therefore cannot reflect a strategic adjustment. Rather, decreased SNR must reflect an unavoidable reaction to the occurrence of negative feedback. Such maladaptive responses are not unprecedented. Several previous studies have reported declines in performance after negative feedback, particularly when it is unexpected (Notebaert et al., 2009; Rabbitt and Rodgers, 1977).

Our results suggest a new interpretation of PES as an attempt to adaptively compensate for a maladaptive consequence of negative feedback. This conjunction explains why RT and accuracy appeared unrelated in many past studies of PES (Danielmeier and Ullsperger, 2011). When the reduction of urgency and sensitivity are balanced, RTs will increase without appreciable changes in accuracy (King et al., 2010; Rabbitt and Rodgers, 1977). When the reduction of urgency (increase of decision bound) overcompensates, PES will be accompanied by increased accuracy, as observed in some studies (Danielmeier and Ullsperger, 2011; Laming, 1979). An imbalance in the opposite direction will result in reduced accuracy, as observed in our post-error speeding subjects and a minority of previous studies (Rabbitt and Rodgers, 1977). We suggest that the key to understanding the diversity of PES behavior is the balance between changes of decision criterion and changes of stimulus sensitivity.

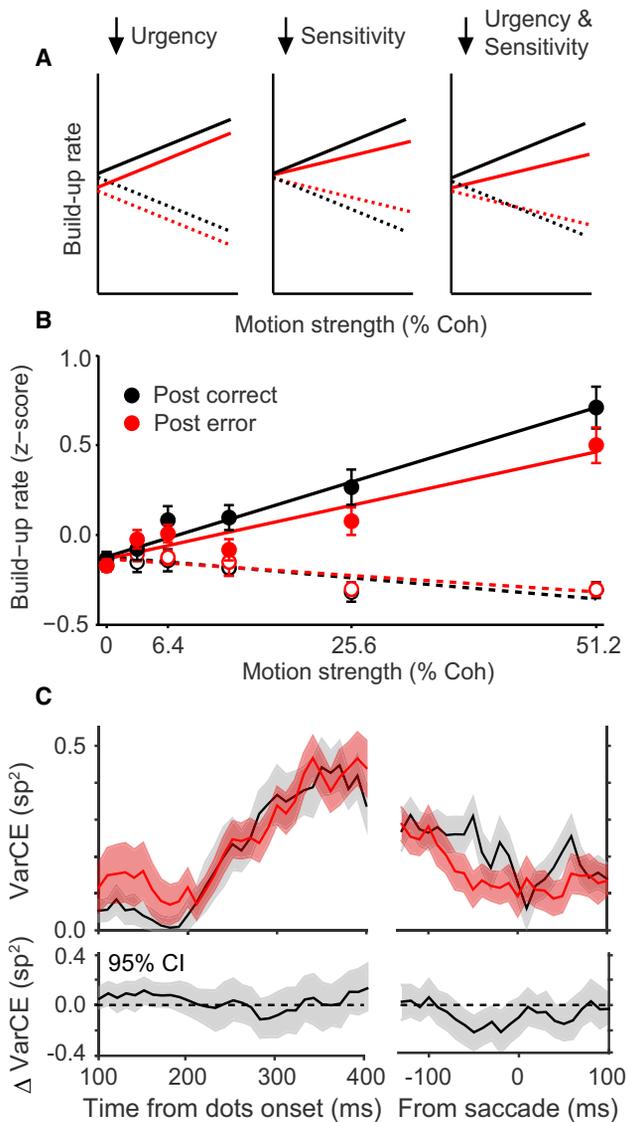


Figure 7. Sensitivity, Not Noise, Changes with PES

(A) Schematic changes in buildup rate under different hypotheses of PES. Reduction of urgency alone (left) reduces buildup rate for all motion strengths. Reduction of sensitivity alone (middle) induces stimulus strength-dependent reduction of buildup rate for T_{in} motions and increase of buildup rate for T_{opp} motions. The combination of reduced urgency and sensitivity (right) predicts a distinct pattern that combines the abovementioned trends.

(B) Measured buildup rates of LIP responses for monkeys O and N during decision formation as a function of motion strength for T_{in} (solid circles) and T_{opp} (open circles) directions. Changes of buildup rates following errors are consistent with a combined change in urgency and sensitivity (A, right). Error bars denote SEM.

(C) Variance of the firing rates that underlie spiking activity is unchanged following errors. This variance indicates a combination of stimulus and neural noise and can be quantified with the VarCE (Equation S4). Top: VarCE for T_{in} for post-correct (black) and post-error (red) trials. Shaded regions indicate $\pm \text{SE}$. Bottom: 95% confidence interval (CI) around the post-error minus post-correct VarCE.

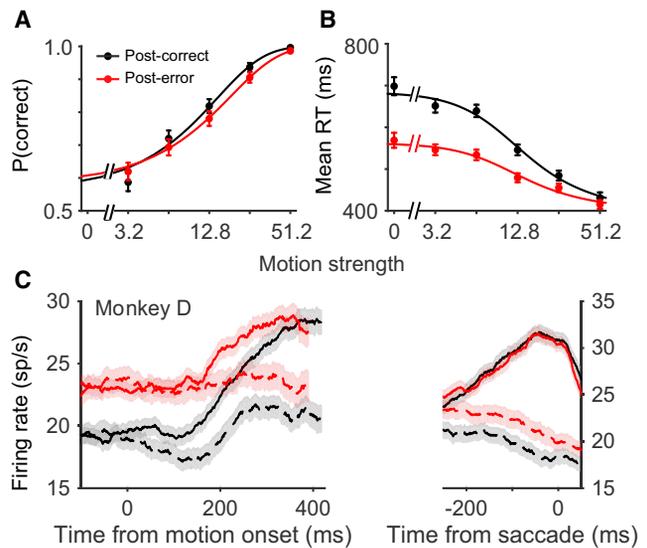


Figure 8. Post-error Speeding Is Associated with a Change in Start Point

(A and B) Psychometric and chronometric functions of subjects that exhibited post-error speeding (monkey D and subject S7). Conventions are similar to those in Figures 1B–1E.

(C) Average responses of the population of LIP neurons recorded from monkey D. The start point of the integration was higher after errors, compatible with a static reduction of effective decision bound height. Conventions are the same as in Figure 5A.

This new hypothesis generates several testable predictions. One prediction is that increasing the delay following feedback facilitates recovery of sensitivity from negative feedback, thereby mitigating the need to slow down. We tested this prediction through a follow-up study in which we manipulated the delay between feedback and the start of the next trial (i.e., the intertrial interval [ITI]; 0.35–3.2 s, see Supplemental Experimental Procedures). We observed that for longer ITIs, PES was reduced (Figure S6), compatible with previous studies (e.g., Danielmeier and Ullsperger, 2011).

The effect of negative feedback on sensitivity also suggests a link between the subject's expectations and PES, leading to a second prediction: unexpected negative feedback should produce larger declines in sensitivity and a larger PES. We tested this prediction by fitting a modified version of the model with an additional parameter that scaled post-error sensitivity by error-trial motion strength. The results indicated a larger decline in sensitivity when subjects received negative feedback in trials with stronger motion (Equation S2; $\beta = -0.86 \pm 0.345$, $p = 0.03$), compatible with our prediction because negative feedback is less expected on those trials (Kiani et al., 2014a). Our prediction is partially consistent with theories that assume error detection and PES are driven by differences between predicted and actual outcomes (Holroyd et al., 2005). However, our results suggest that this component of PES is not necessarily adaptive. Distinguishing whether expectations drive adaptive or maladaptive components of PES could be addressed through experiments that explicitly probe subjects' expectations before feedback delivery (Purcell and Kiani, 2015, Soc. Neurosci., conference).

Neural Mechanisms of PES

A complete understanding of PES requires us to understand (1) how errors are detected and (2) how they influence future sensory processing and decision strategies. The role of medial frontal cortex in error detection is well known. Medial frontal neurons discharge immediately after erroneous responses or feedback (Purcell et al., 2012; Schall et al., 2002), contributing to error-related negativity in extracranial electroencephalographic signal (Gehring et al., 1993) and BOLD responses associated with performance monitoring (King et al., 2010). Further, artificial inactivation of medial frontal cortex eliminates PES in rodents (Narayanan et al., 2013), and medial frontal lesions in humans reduce or eliminate PES (di Pellegrino et al., 2007), indicating a critical connection between performance monitoring processes and PES.

Studies on medial frontal cortex, however, do not tell us how errors affect sensory processing and decision strategies. Previous studies have reported correlations between medial frontal cortex and decision-related frontal and parietal areas during PES (King et al., 2010), but the nature of this interaction is unclear at a single-neuron level. Anatomically, medial frontal cortex is not well positioned to accumulate sensory evidence, because it receives relatively sparse inputs from sensory areas (Huerta and Kaas, 1990). During perceptual decision making, medial frontal neurons may not exhibit choice-selective signals (Purcell et al., 2012), and firing rates do not converge at a fixed bound prior to response execution (Stuphorn et al., 2010). In contrast, LIP is a node in a larger network of brain areas involved in integration of sensory inputs to plan eye movements. In the direction discrimination task, LIP neurons represent integration of sensory evidence to a decision bound (Gold and Shadlen, 2007). Our recordings from LIP provide crucial new insights into the mechanisms by which errors modify future decisions.

Neural responses that provide access to the dynamics of the integration process are key to distinguish hypotheses that could not be differentiated on the basis of behavior alone. Below, we expand on three components of the decision-making process that have been implicated in PES by past studies.

Delays in Decision Onset

Our observation that the onset of evidence accumulation is indistinguishable for post-correct and post-error trials is inconsistent with theories of PES that hypothesize delays in the start of evidence accumulation following errors (Laming, 1979; Rabbitt and Rodgers, 1977). We cannot directly rule out delayed activation of motor processes downstream of LIP; for example, slowing following countermanded saccades has been shown to produce delayed activation in oculomotor neurons (Pouget et al., 2011). However, the non-uniform slowing across RTs and absence of increased non-decision time in the model indicate that delayed motor activation is not sufficient to explain PES and is unlikely to be a core component of increased RTs.

Quality of Evidence and Efficiency of Accumulation

Although the DDM indicated that reduction of SNR was a core component of PES, the model could not separately quantify the role of sensitivity and noise. The inadequacy of DDM is an inherent feature shared by all bounded accumulation models. To provide intuition about this inadequacy, we note that fitting a model in which bound, sensitivity, and noise are all free to

vary does not generate a unique solution. Broadly speaking, to achieve a unique fit, one must fix at least one of the three factors and express the remaining two in the units of the fixed one. It is therefore implausible to use the pattern of choices and RTs to independently assess post-error changes of sensitivity and noise in the presence of bound changes.

LIP responses addressed this conundrum by showing no change in VarCE following errors. VarCE reflects the sum of sensory and accumulation noise (Churchland et al., 2011), and its constancy suggests that elevated neural noise is not a major contributor to PES. On the other hand, a reduction of sensitivity shows clearly in the buildup rates of LIP neurons. In our task, sensory evidence for direction discrimination is provided by extra-striate visual cortical areas, MT and MST (Britten et al., 1992; Ditterich et al., 2003; Fetsch et al., 2014). Reduction of sensitivity in parietal neurons may reflect a reduction in the quality of the sensory representation in MT and MST (Treue and Martínez Trujillo, 1999) or the readout of sensory information by the decision-making network (Green et al., 2010). On the basis of past fMRI studies, we speculate that the former is less likely (Danielmeier and Ullsperger, 2011; King et al., 2010), but distinguishing these two possibilities requires electrophysiological recordings from MT and MST.

Decision Bound

Bound adjustments in PES are likely implemented through modulation of a dynamic urgency signal, an evidence-independent input that equally drives all choice-selective neural populations (Churchland et al., 2008; Hanks et al., 2014). Decreased urgency causes neural responses to rise more slowly, increasing the evidence required to reach the bound. Similar changes in LIP and prefrontal neurons have been shown to implement changes of decision bound for speed-accuracy trade-off (Hanks et al., 2014; Heitz and Schall, 2012). Thus, our results contribute to mounting evidence that changes in decision bound are not implemented through elevated endpoint, as one might expect, but instead through mechanisms that directly accelerate or decelerate accumulation of evidence toward a fixed threshold (Hanks et al., 2014; Lo et al., 2015).

Post-error reduction of urgency increases the accumulated evidence for the decision and neutralizes the reduction of SNR. In theory, such an effective increase of decision bound could also have been implemented by a static reduction of the start point of the accumulation process. There is, however, a limitation to this alternative. Because firing rates are bounded at zero, decreasing the start point has an inherently restricted range. A similar limitation, however, does not apply if one's goal is to reduce the effective bound height to respond faster following errors, as did monkey D in our experiment (Figure 8).

A dynamic modulation of decision bound may be computationally advantageous, compared with a static change of the bound. Note that the difference of post-correct and post-error urgency signals grows with time (Figure 6). The effective increase in the decision bound is therefore largest for longer RTs, which are associated with intermediate and weak motion stimuli (Figure 1). For the same stimuli, accuracy is most susceptible to a reduction of SNR and also can improve the most with an increase of decision bound. In contrast, for the strongest stimuli, accuracy tends to be high and is unlikely to significantly benefit

from increasing the decision bound. Manipulating decision bound through urgency ensures the increase of decision time happens where it matters the most.

The origin of urgency signal is unknown, but it may be shaped by the cost of acquiring evidence (Drugowitsch et al., 2012) and implemented by mechanisms that encode elapsed time (Jazayeri and Shadlen, 2015). The post-error changes of urgency, however, are likely to originate from structures involved in error detection (Purcell et al., 2012; Schall et al., 2002). Although it is tempting to conclude that these error-related signals directly furnish changes of decision bound in the following trials, it is important to note that these signals are limited to the period immediately after the error and rarely overlap with the following decisions. Most likely, the effect on urgency is mediated through mechanisms that store long-term history of past behavior (e.g., Kiani et al., 2015; Tsujimoto et al., 2010). Future studies will shed light on this issue.

The changes of urgency in LIP responses can be part of a more widespread effect in the decision-making and action-planning network. For example, pre-supplementary motor area is inhibited (Danielmeier and Ullsperger, 2011; King et al., 2010) and corticospinal excitability is reduced in PES (Amengual et al., 2013). Such inhibitions, if also present in our task, are unlikely to be a purely motor (i.e., post-decisional) effect because they would predict equal elongation of RT for all stimulus strengths, incompatible with our experimental observations (Figures 1, 3, and S1). Rather, they are likely to be part of the decision-making process itself. Mounting evidence suggests that the decision-making process is distributed in a network that includes motor areas (Cisek and Kalaska, 2010; Gold and Shadlen, 2000; Hanes and Schall, 1996; Noorbalooshi et al., 2015; Ratcliff et al., 2003; Selen et al., 2012). From this perspective, we view our parietal recordings as a window onto the activity of this distributed network, and so the reductions in motor activity likely reflect another manifestation of the reduced urgency that we observe in our parietal neurons.

Conclusions

The neural mechanisms of PES have been hotly debated, but different hypotheses have been difficult to distinguish using only behavioral data or noninvasive methods for monitoring neural activity. Using quantitative modeling and electrophysiology, we provide a new account of PES as an active reduction of response urgency to balance reductions of sensitivity and maintain performance. This dual mechanism quantitatively explains post-error behavior and offers a parsimonious description for its diversity in past studies.

EXPERIMENTAL PROCEDURES

Behavioral Task

Seven humans (four female, three male; ages 18–30 years) and four rhesus monkeys (two male, two female; ages 4–12 years) were trained to perform a motion direction discrimination task. All subjects were naive to the purpose of the experiment. Informed written consent was obtained from all human subjects. Experimental procedures for humans were approved by the Institutional Review Board at New York University. Experimental procedures for monkeys conformed to the National Institutes of Health *Guide for the Care and Use of Laboratory Animals* and were approved by the animal care committees at

the University of Washington (monkeys N and B) and Stanford University (monkeys O and D). Data from monkeys N and B have been previously published (Roitman and Shadlen, 2002).

Subjects were seated in a semi-dark room in front of a cathode ray tube monitor (frame rate, 75 Hz) with their heads stabilized (surgically implanted head post for monkeys, chin and forehead support for humans). They began each trial by fixating a small red point at the center of the screen (fixation point [FP], 0.3° diameter). After a variable delay (truncated exponential distribution; mean = 300–700 ms), two red targets (0.5°) appeared on opposite sides of the screen equidistant from the FP. After another random delay, the dynamic random-dots stimulus (Britten et al., 1992) appeared within a 5°–7° circular aperture centered on the FP. The stimulus consisted of three independent sets of moving dots shown in consecutive frames. Each set of dots was shown for one video frame and then replotted three frames later ($\Delta t = 40$ ms; density, 16.7 dots/deg²/s). When replotted, a subset of dots were offset from their original location (speed, 5 deg/s) while the remaining dots were placed randomly. The percentage of coherently displaced dots determined the strength of motion. Subjects reported their perceived direction of motion with a saccadic eye movement to the choice target in the direction of motion. Subjects were free to indicate their choice any time after motion onset. RT was defined as the difference between saccade initiation time and motion onset. See [Supplemental Experimental Procedures](#) for additional task details.

Behavioral Data Analyses

For all analyses, “post-error” refers to trials following an error feedback, and “post-correct” refers to trials that followed a correct feedback. In addition, we used only post-correct trials that are also pre-error to guard against slow fluctuations in behavioral state (see [Supplemental Experimental Procedures](#)).

We used a logistic regression to assess the effects of motion strength and errors on accuracy on the subsequent trial:

$$\text{logit}[P(\text{correct})] = \beta_1 C + \beta_2 C_l, \quad (\text{Equation 1})$$

where $\text{logit}(p)$ is $\log(p/1-p)$, C is the motion strength on the current trial, l indicates the outcome of the previous trial (0 = correct, 1 = error), and β_i are the regression coefficients. β_1 tests for the effect of motion strength on choice, and β_2 tests for the effect of previous-trial outcome on the slope of the psychometric function.

Similarly, we assessed the effects of motion strength and errors on RT with the following linear regression:

$$T = \beta_0 + \beta_1 C + \beta_2 l + \beta_3 C_l, \quad (\text{Equation 2})$$

where T is the RT on the current trial. We obtained identical results using a nonlinear regression analysis (hyperbolic tangent).

We tested whether PES was greatest at shorter or longer RTs using a linear regression model of the following form:

$$\Delta T = \beta_0 + \beta_1 C + \beta_2 D, \quad (\text{Equation 3})$$

where ΔT indicates the mean RT difference between post-error and post-correct trials for coherence, C , and RT decile, D . Deciles were computed separately for each coherence and trial type (post-correct or post-error). Positive β_2 indicates maximal slowing at longer RTs ($H_0: \beta_2 = 0$). We obtained similar results using different numbers of quantiles.

DDM

We fit a DDM to each subject's behavior following correct and error trials to evaluate alternative hypotheses about the mechanisms responsible for PES. Choice and RT result from integration of noisy sensory evidence over time to a decision criterion or bound. The model parameters are as follows. The sensitivity parameter, k , determines the linear scaling of the mean rate of accumulation (i.e., drift rate) with motion strength (C). Therefore, $\mu = kC$ is the mean of the distribution of momentary evidence per time unit. The SD of the momentary evidence, σ , is fixed to 1 in the majority of model fits. For one variation of the model, we compared the standard model with the effect of varying SNR of sensory evidence by changing the noise (σ) and fixing k . The non-decision time parameter, t_0 , determines the mean time necessary for sensory processing and motor delays before and after the decision process. We further assume

that non-decision time is normally distributed with SD, s_{t_0} . The bound height, $B(t)$, determines the amount of evidence that must be accumulated for a choice at decision time t . We modeled evidence-independent response urgency as a collapsing bound over time (Hanks et al., 2014). A hyperbolic function was chosen for the shape of the bound collapse (Churchland et al., 2008):

$$|B(t)| = b - u_{\infty} \frac{t}{t + \tau_{1/2}}, \quad (\text{Equation 4})$$

where b is the initial bound height, u_{∞} is the asymptotic reduction in bound height, and $\tau_{1/2}$ is the time to reach 50% reduction. Similar results were obtained with other functional forms for urgency. Finally, an offset in the start point of the accumulation, z , and a bias in the duration of non-decision time for right/left choices, t_{0bias} , accounted for choice biases and choice-dependent RT difference that are occasionally observed in both human and non-human primates. See [Supplemental Experimental Procedures](#) for details of model fitting. Briefly, we fit the full RT distributions, irrespective of choice, using a maximum likelihood procedure. In most fits, we allowed four model parameters (k , t_0 , u_{∞} , and $\tau_{1/2}$) to vary between post-error and post-correct conditions.

Neural Analyses

Single-neuron responses were recorded from area LIP of four rhesus monkeys while they performed the motion discrimination task (see [Supplemental Experimental Procedures](#)). Peristimulus time histograms (PSTHs) were created by averaging neural responses aligned to different task events. For illustration purposes, they were smoothed by convolution with a 100 ms boxcar filter. We divided trials according to whether the chosen target was inside (T_{in}) or opposite (T_{opp}) the RF of the neuron and whether the trial was post-correct or post-error, as defined above.

We estimated the start time of evidence integration on the basis of PSTHs of trials ending with T_{in} saccades. A 40 ms window moved backward along the PSTH from the saccade onset toward motion onset in 1 ms steps. The onset of integration was defined as the first time when (1) activity made a transition from a ramp to a stable state (ramp was defined with a significant Spearman correlation for 10 consecutive steps, $p < 0.01$) and (2) firing rate showed at least a 20% drop compared with 100 ms before the saccade onset (Pouget et al., 2011). Other methods for measuring activity onset (e.g., dip time, see below) produced similar conclusions. Also, similar results were obtained for individual neurons and the population.

The firing rates associated with the beginning and end of the evidence accumulation process were measured in two distinct time windows. For the start point firing rate, we used a 50ms window centered on the post-stimulus dip in LIP activity. The dip is hypothesized to reflect a reset of the integration process and happens 50–150 ms after motion onset. We defined the dip as the minimum firing rate for each neuron in that period (mean dip time = 107 ± 3.34 ms). The dip time closely corresponds with the accumulation onset time estimated above. Thus, similar results were obtained by using the firing rate around the onset time. Using other measurements of start point firing rate (e.g., the baseline period 0–200 ms before motion onset) led to identical conclusions. We estimated the endpoint of accumulation as the mean firing rate 75–25 ms before the saccade onset, a time period at which firing rates converge to a common level irrespective of RT and motion strength (Figure S3; Churchland et al., 2008; Roitman and Shadlen, 2002). Only trials in which the T_{in} target was chosen were included in this analysis. We also tested for differences in the overall excursion, that is, the difference between the start point and endpoint firing rate of each neuron. We assessed the significance of activity changes in each epoch using a mixed-measures ANOVA with motion strength and previous trial outcome as fixed effects and cell identity as a random effect.

We quantified the urgency signal using two different methods: averaging neural responses across all motion directions, motion strength, and choices (Churchland et al., 2008), and piecewise estimation of the derivative of the urgency signal (Hanks et al., 2011). See [Supplemental Experimental Procedures](#) for details. Parameterization of post-correct and post-error urgency signals was done by fitting a hyperbolic function (Equation S3) in the epoch from the response dip until 450 ms after motion onset (Figure 6A).

We computed the buildup rate of LIP neurons on a trial-by-trial basis by convolving the spike train with a Gaussian kernel (SD = 25 ms) to estimate the smoothed rate function and then by applying a linear regression to this rate function. The analysis was performed on an epoch starting at the characteristic dip and ending 100 ms before saccade initiation. To combine the buildup rates across neurons, we standardized them for each neuron (Z score). We tested for post-error changes of buildup rate using the following linear regression:

$$v = \beta_0 + \beta_1 I + \beta_2 C J_1 + \beta_3 C J_2 + \beta_4 C J_1 I + \beta_5 C J_2 I, \quad (\text{Equation 5})$$

where v indicates the normalized buildup rate and C is the motion strength on each trial. I is an indicator variable for the previous trial outcome. J_1 and J_2 are indicator variables for motion direction: J_1 equals 1 for motion toward the neuron's RF and 0 otherwise; J_2 equals 1 for motion away from RF and 0 otherwise. β_2 and β_3 indicate the change of buildup rate with motion strength toward T_{in} and T_{opp} (i.e., the neural sensitivity), and β_4 and β_5 indicate the change of sensitivity following errors. The null hypothesis is no change in sensitivity (H_0 : $\beta_4 = 0$ and $\beta_5 = 0$).

We also tested whether the variation in buildup rate following errors can be entirely accounted for by changes in RTs. First, we removed the covariation with RT using a regression model:

$$v = \beta_0 + \beta_1 T. \quad (\text{Equation 6})$$

Then the residual buildup rate that could not be explained by RT variations ($v_r = v - \beta_0 - \beta_1 T$) was used in Equation 5.

To test for changes in neural variance during PES, we measured the across-trial firing rate variability as the VarCE (Churchland et al., 2011). See [Supplemental Experimental Procedures](#) for details. Briefly, the analysis assumes that the spike count in each epoch is a stochastic realization of a point process governed by a rate parameter, which itself is randomly drawn from a distribution. VarCE estimates the variance of this rate distribution at each moment by subtracting from the total spike count variance an estimate of the point-process variance.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, six figures, and two tables and can be found with this article online at <http://dx.doi.org/10.1016/j.neuron.2015.12.027>.

AUTHOR CONTRIBUTIONS

B.A.P. and R.K. conceived and designed the experiment, collected data, performed the analyses, and wrote the paper.

ACKNOWLEDGMENTS

This work was supported by a Sloan Research Fellowship, a Young Investigator Grant from the Brain & Behavior Research Foundation, and a Whitehall Research Grant to R.K. B.A.P. was supported by National Institutes of Health training grant T32EY007136 and a post-doctoral fellowship from the Simons Collaboration on the Global Brain. We thank Paul Glimcher, Michael Landy, Fred Rieke, and Paul Miller for suggesting that we explore models that assume variation in decision noise. We thank Long Sha, Christina Hatch, and Robbe Goris for helpful discussions and Saleh Esteki for assistance with human data collection.

Received: June 26, 2015

Revised: September 21, 2015

Accepted: December 15, 2015

Published: January 21, 2016

REFERENCES

Amengual, J.L., Marco-Pallarés, J., Richter, L., Oung, S., Schweikard, A., Mohammadi, B., Rodríguez-Fornells, A., and Münte, T.F. (2013). Tracking

- post-error adaptation in the motor system by transcranial magnetic stimulation. *Neuroscience* 250, 342–351.
- Bogacz, R., Brown, E., Moehlis, J., Holmes, P., and Cohen, J.D. (2006). The physics of optimal decision making: a formal analysis of models of performance in two-alternative forced-choice tasks. *Psychol. Rev.* 113, 700–765.
- Botvinick, M.M., Braver, T.S., Barch, D.M., Carter, C.S., and Cohen, J.D. (2001). Conflict monitoring and cognitive control. *Psychol. Rev.* 108, 624–652.
- Britten, K.H., Shadlen, M.N., Newsome, W.T., and Movshon, J.A. (1992). The analysis of visual motion: a comparison of neuronal and psychophysical performance. *J. Neurosci.* 12, 4745–4765.
- Busemeyer, J.R., and Townsend, J.T. (1993). Decision field theory: a dynamic-cognitive approach to decision making in an uncertain environment. *Psychol. Rev.* 100, 432–459.
- Churchland, A.K., Kiani, R., and Shadlen, M.N. (2008). Decision-making with multiple alternatives. *Nat. Neurosci.* 11, 693–702.
- Churchland, A.K., Kiani, R., Chaudhuri, R., Wang, X.J., Pouget, A., and Shadlen, M.N. (2011). Variance as a signature of neural computations during decision making. *Neuron* 69, 818–831.
- Cisek, P., and Kalaska, J.F. (2010). Neural mechanisms for interacting with a world full of action choices. *Annu. Rev. Neurosci.* 33, 269–298.
- Danielmeier, C., and Ullsperger, M. (2011). Post-error adjustments. *Front. Psychol.* 2, 233.
- di Pellegrino, G., Ciaramelli, E., and Làdavas, E. (2007). The regulation of cognitive control following rostral anterior cingulate cortex lesion in humans. *J. Cogn. Neurosci.* 19, 275–286.
- Ditterich, J., Mazurek, M.E., and Shadlen, M.N. (2003). Microstimulation of visual cortex affects the speed of perceptual decisions. *Nat. Neurosci.* 6, 891–898.
- Drugowitsch, J., Moreno-Bote, R., Churchland, A.K., Shadlen, M.N., and Pouget, A. (2012). The cost of accumulating evidence in perceptual decision making. *J. Neurosci.* 32, 3612–3628.
- Dutilh, G., Forstmann, B.U., Vandekerckhove, J., and Wagenmakers, E.J. (2013). A diffusion model account of age differences in posterror slowing. *Psychol. Aging* 28, 64–76.
- Fetsch, C.R., Kiani, R., Newsome, W.T., and Shadlen, M.N. (2014). Effects of cortical microstimulation on confidence in a perceptual decision. *Neuron* 83, 797–804.
- Gehring, W.J., Goss, B., Coles, M.G., Meyer, D.E., and Donchin, E. (1993). A neural system for error detection and compensation. *Psychol. Sci.* 4, 385–390.
- Gold, J.I., and Shadlen, M.N. (2000). Representation of a perceptual decision in developing oculomotor commands. *Nature* 404, 390–394.
- Gold, J.I., and Shadlen, M.N. (2007). The neural basis of decision making. *Annu. Rev. Neurosci.* 30, 535–574.
- Goldfarb, S., Wong-Lin, K., Schwemmer, M., Leonard, N.E., and Holmes, P. (2012). Can post-error dynamics explain sequential reaction time patterns? *Front. Psychol.* 3, 213.
- Green, C.S., Pouget, A., and Bavelier, D. (2010). Improved probabilistic inference as a general learning mechanism with action video games. *Curr. Biol.* 20, 1573–1579.
- Hanes, D.P., and Schall, J.D. (1996). Neural control of voluntary movement initiation. *Science* 274, 427–430.
- Hanks, T.D., Mazurek, M.E., Kiani, R., Hopp, E., and Shadlen, M.N. (2011). Elapsed decision time affects the weighting of prior probability in a perceptual decision task. *J. Neurosci.* 31, 6339–6352.
- Hanks, T., Kiani, R., and Shadlen, M.N. (2014). A neural mechanism of speed-accuracy tradeoff in macaque area LIP. *eLife* 3, 3.
- Heitz, R.P., and Schall, J.D. (2012). Neural mechanisms of speed-accuracy tradeoff. *Neuron* 76, 616–628.
- Holroyd, C.B., Yeung, N., Coles, M.G., and Cohen, J.D. (2005). A mechanism for error detection in speeded response time tasks. *J. Exp. Psychol. Gen.* 134, 163–191.
- Huerta, M.F., and Kaas, J.H. (1990). Supplementary eye field as defined by intracortical microstimulation: connections in macaques. *J. Comp. Neurol.* 293, 299–330.
- Jazayeri, M., and Shadlen, M.N. (2015). A Neural Mechanism for Sensing and Reproducing a Time Interval. *Curr. Biol.* 25, 2599–2609.
- Kiani, R., Corthell, L., and Shadlen, M.N. (2014a). Choice certainty is informed by both evidence and decision time. *Neuron* 84, 1329–1342.
- Kiani, R., Cueva, C.J., Reppas, J.B., and Newsome, W.T. (2014b). Dynamics of neural population responses in prefrontal cortex indicate changes of mind on single trials. *Curr. Biol.* 24, 1542–1547.
- Kiani, R., Cueva, C.J., Reppas, J.B., Peixoto, D., Ryu, S.I., and Newsome, W.T. (2015). Natural grouping of neural responses reveals spatially segregated clusters in prearcuate cortex. *Neuron* 85, 1359–1373.
- King, J.A., Korb, F.M., von Cramon, D.Y., and Ullsperger, M. (2010). Post-error behavioral adjustments are facilitated by activation and suppression of task-relevant and task-irrelevant information processing. *J. Neurosci.* 30, 12759–12769.
- Laming, D. (1979). Choice reaction performance following an error. *Acta Psychol. (Amst.)* 43, 199–224.
- Link, S.W. (1992). *The Wave Theory of Difference and Similarity* (Lawrence Erlbaum Associates).
- Lo, C.C., Wang, C.T., and Wang, X.J. (2015). Speed-accuracy tradeoff by a control signal with balanced excitation and inhibition. *J. Neurophysiol.* 114, 650–661.
- Narayanan, N.S., Cavanagh, J.F., Frank, M.J., and Laubach, M. (2013). Common medial frontal mechanisms of adaptive control in humans and rodents. *Nat. Neurosci.* 16, 1888–1895.
- Noorbaloochi, S., Sharon, D., and McClelland, J.L. (2015). Payoff Information Biases a Fast Guess Process in Perceptual Decision Making under Deadline Pressure: Evidence from Behavior, Evoked Potentials, and Quantitative Model Comparison. *J. Neurosci.* 35, 10989–11011.
- Notebaert, W., Houtman, F., Opstal, F.V., Gevers, W., Fias, W., and Verguts, T. (2009). Post-error slowing: an orienting account. *Cognition* 111, 275–279.
- Pouget, P., Logan, G.D., Palmeri, T.J., Boucher, L., Paré, M., and Schall, J.D. (2011). Neural basis of adaptive response time adjustment during saccade countermanding. *J. Neurosci.* 31, 12604–12612.
- Purcell, B.A., Weigand, P.K., and Schall, J.D. (2012). Supplementary eye field during visual search: salience, cognitive control, and performance monitoring. *J. Neurosci.* 32, 10273–10285.
- Rabbitt, P., and Rodgers, B. (1977). What does a man do after he makes an error? An analysis of response programming. *Q. J. Exp. Psychol.* 29, 727–743.
- Ratcliff, R., Cherian, A., and Segraves, M. (2003). A comparison of macaque behavior and superior colliculus neuronal activity to predictions from models of two-choice decisions. *J. Neurophysiol.* 90, 1392–1407.
- Roitman, J.D., and Shadlen, M.N. (2002). Response of neurons in the lateral intraparietal area during a combined visual discrimination reaction time task. *J. Neurosci.* 22, 9475–9489.
- Schall, J.D., Stuphorn, V., and Brown, J.W. (2002). Monitoring and control of action by the frontal lobes. *Neuron* 36, 309–322.
- Schroder, H.S., and Moser, J.S. (2014). Improving the study of error monitoring with consideration of behavioral performance measures. *Front. Hum. Neurosci.* 8, 178.
- Selen, L.P., Shadlen, M.N., and Wolpert, D.M. (2012). Deliberation in the motor system: reflex gains track evolving evidence leading to a decision. *J. Neurosci.* 32, 2276–2286.
- Sergeant, J.A., and van der Meere, J. (1988). What happens after a hyperactive child commits an error? *Psychiatry Res.* 24, 157–164.
- Standage, D., You, H., Wang, D.H., and Dorris, M.C. (2011). Gain modulation by an urgency signal controls the speed-accuracy trade-off in a network model of a cortical decision circuit. *Front. Comput. Neurosci.* 5, 7.

Stuphorn, V., Brown, J.W., and Schall, J.D. (2010). Role of supplementary eye field in saccade initiation: executive, not direct, control. *J. Neurophysiol.* *103*, 801–816.

Thura, D., Cos, I., Trung, J., and Cisek, P. (2014). Context-dependent urgency influences speed-accuracy trade-offs in decision-making and movement execution. *J. Neurosci.* *34*, 16442–16454.

Treue, S., and Martínez Trujillo, J.C. (1999). Feature-based attention influences motion processing gain in macaque visual cortex. *Nature* *399*, 575–579.

Tsujimoto, S., Genovesio, A., and Wise, S.P. (2010). Evaluating self-generated decisions in frontal pole cortex of monkeys. *Nat. Neurosci.* *13*, 120–126.

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

Neuron

Supplemental Information

Neural Mechanisms of Post-error Adjustments of Decision Policy in Parietal Cortex

Braden A. Purcell and Roozbeh Kiani

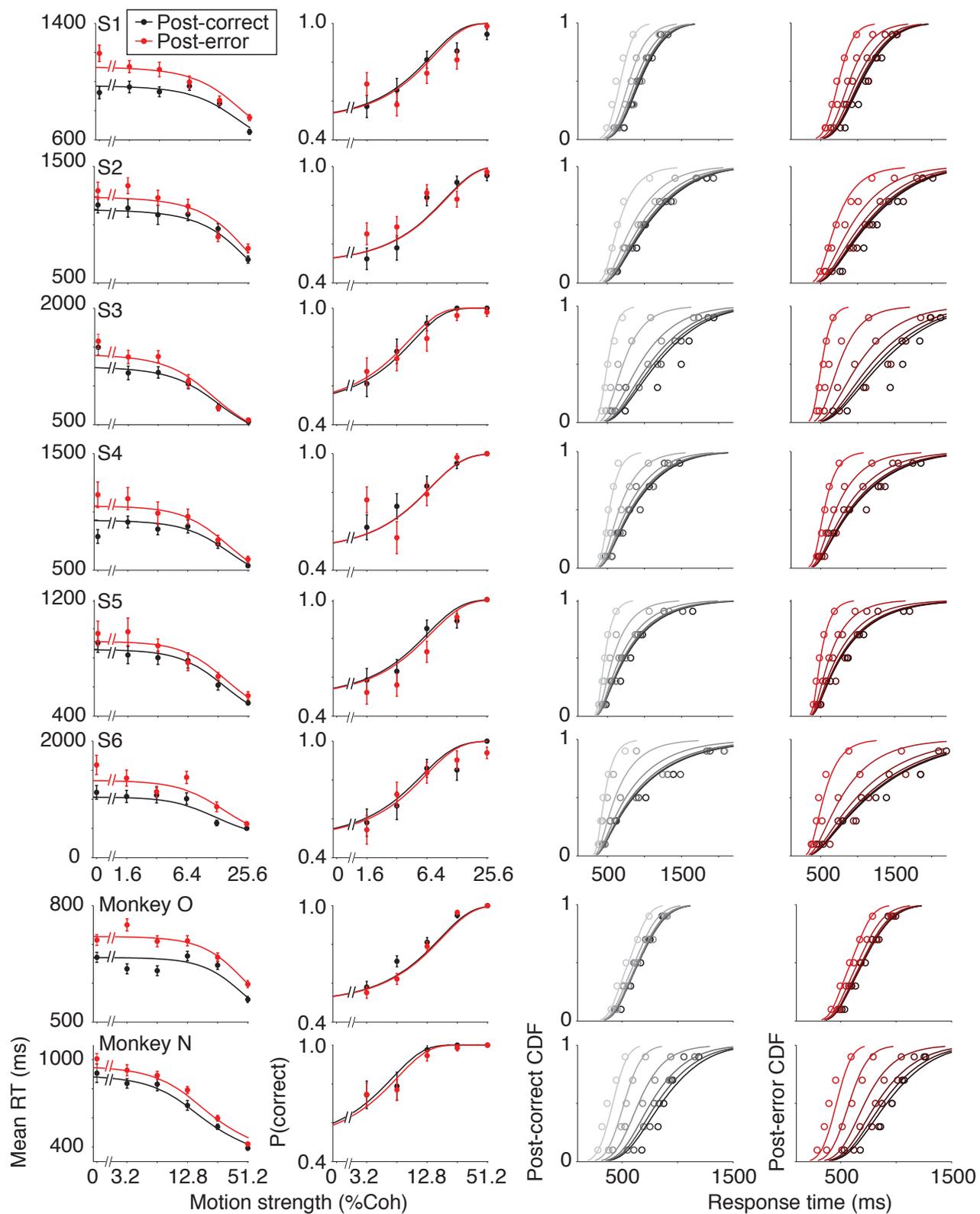


Figure S1. Related to Figure 3. Drift-diffusion model fits and predictions for individual subjects that exhibited post-error slowing. Each row shows fitted chronometric functions (column 1), predicted psychometric functions (column 2), and fitted post-correct (column 3) and post-error (column 4) cumulative RT distributions (CDFs) for each subject. Lines are model fits (columns 1, 3, and 4) or predictions (column 2) and circles are data. Black indicates post-correct trials and red indicates post-error trials. For RT distributions, the brightness increases with motion strength (%Coh).

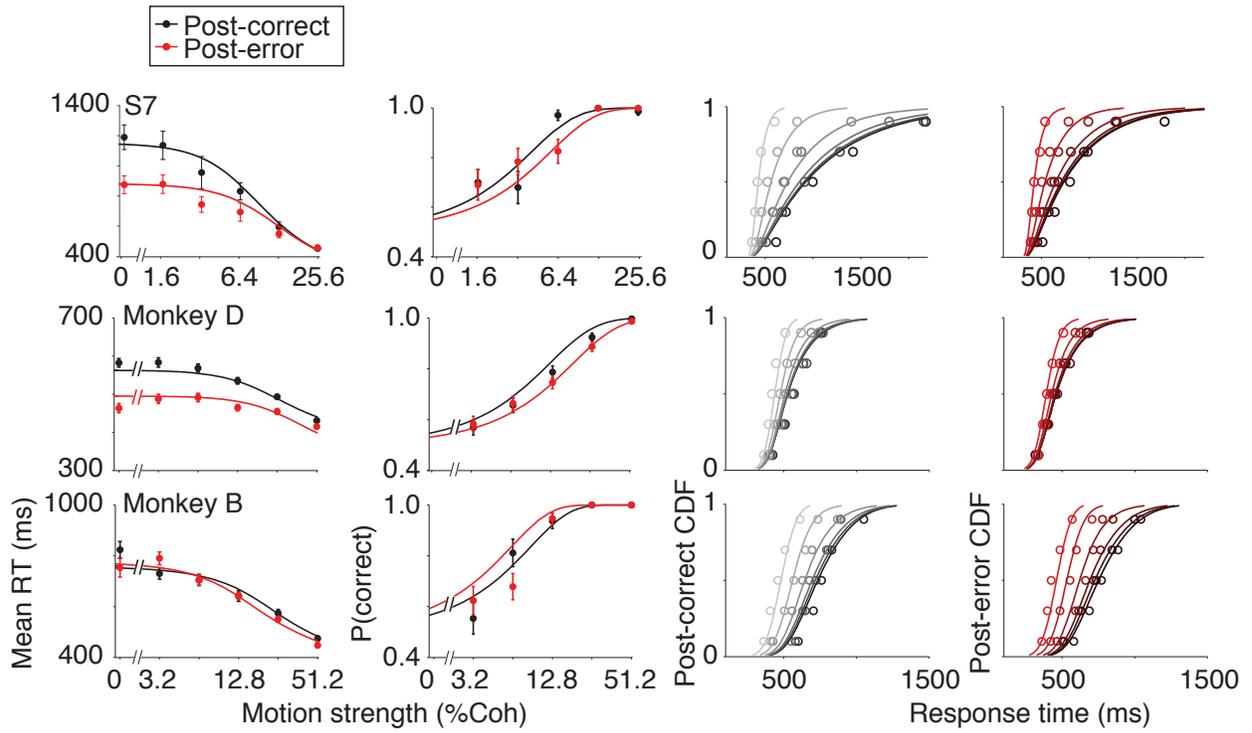


Figure S2. Related to Figures 3, 8, and S5. Drift-diffusion model fits and predictions for individual subjects that did not slow after errors. Conventions are the same as in Figure S1. Human S7 and Monkey D showed post-error speeding. Monkey B showed no RT change following errors.

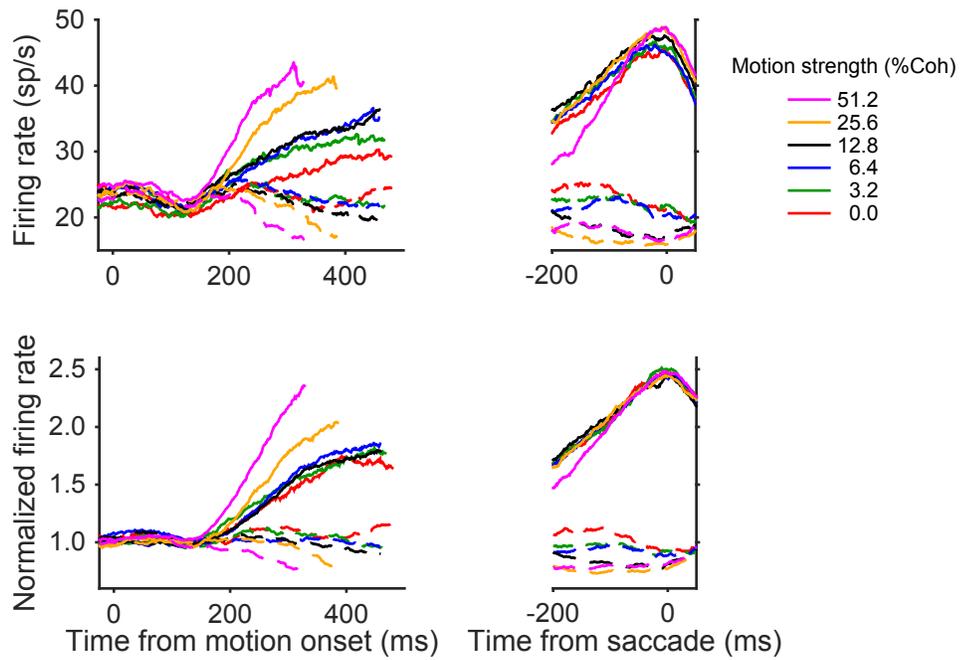


Figure S3, related to Figure 5. LIP responses represent evidence accumulation. Average (top) and normalized (bottom) LIP responses from Monkeys O and N resemble evidence accumulation to a bound. When aligned on motion onset (left), population responses ramp at a rate that is proportional to the motion strength. When aligned on saccade, population responses converge to a fixed level ~70 ms prior to saccade initiation.

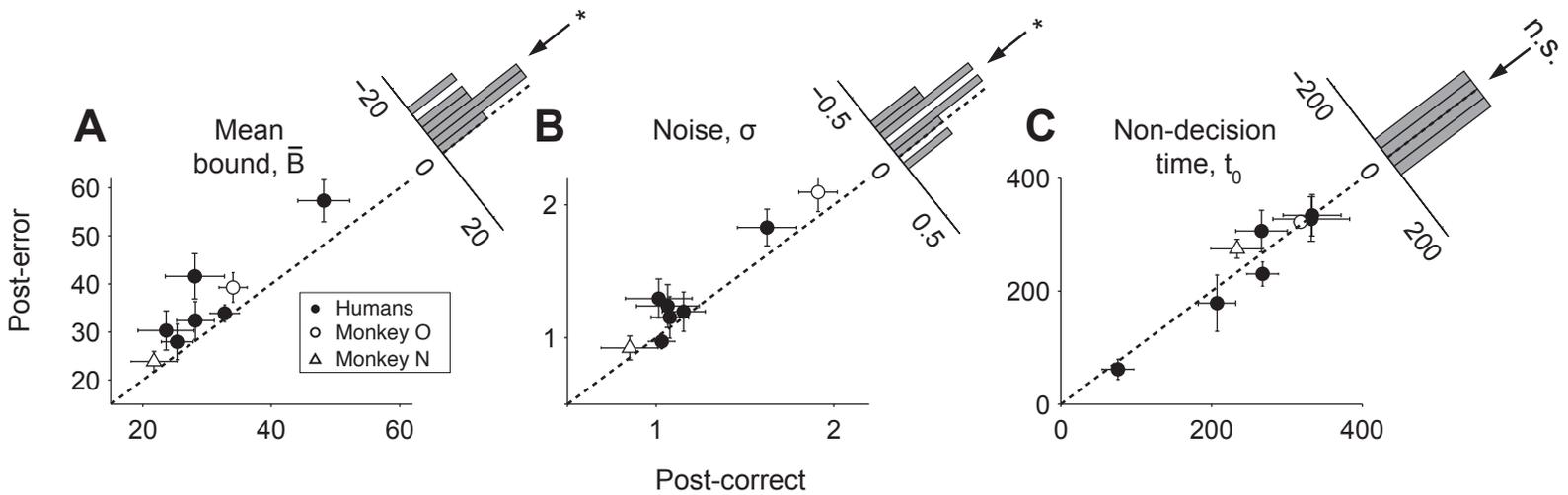


Figure S4, related to Figure 7. Best fitting parameters for a drift-diffusion model in which decision noise varies between post-correct and post-error trials

We fit an alternative formulation of the drift-diffusion model in which sensitivity was fixed and decision noise was free to vary with PES. The best fitting parameters revealed a consistent decrease in urgency resulting in a consistent elevation of decision bound (A), a consistent increase in decision noise (B), but no systematic pattern of changes in non-decision time (C). The quality of fit to behavior was indistinguishable from our main model in which sensitivity and bound varied. However, this alternative model is ruled out based on the absence of changes in neural response variance following errors (see Figure 7C). Conventions are similar to Figure 4C-E.

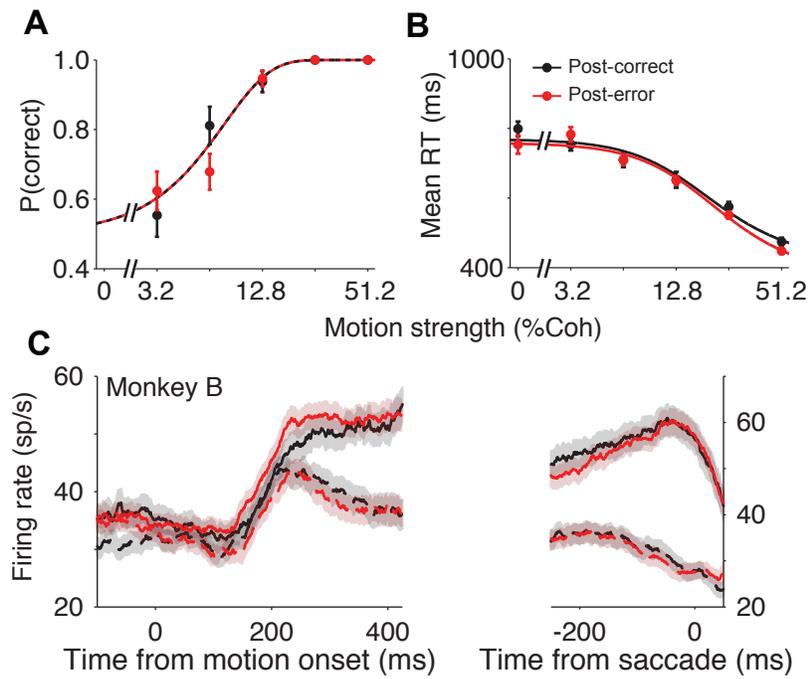


Figure S5, related to Figure 8. Post-error changes in neuronal responses are specific to post-error behavioral adjustments

(A-B) Psychometric and chronometric functions of monkey B that exhibited no post-error adjustment in behavior. Conventions are similar to Figure 1B-E.

(C) Average responses of the population of LIP neurons recorded from monkey B. Responses are identical regardless of the previous trial's outcome.

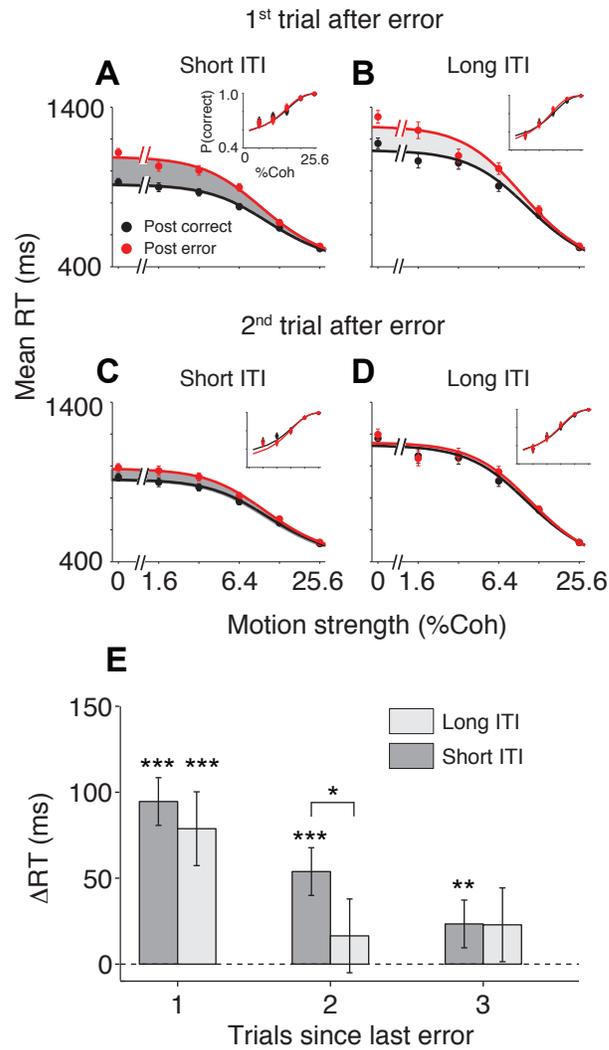


Figure S6, related to Figures 1 and 3. Increasing inter-trial interval (ITI) reduces PES. (A-B) On the first trial after the error, subjects exhibited PES for the tested ITIs (short ITI: 0.35-0.8 s, $\Delta RT = 94 \pm 13.9$ ms; long ITI: 2.4-3.2 s, $\Delta RT = 79 \pm 21.5$ ms). Mean RTs were longer on post-error trials (red) than post-correct (black) trials. The shaded area emphasizes the magnitude of PES (post-error RTs minus post-correct RT), demonstrating that our primary results are not a consequence of selecting a specific ITI. Accuracy was unchanged (insets). (C-D) Two trials after the error, subjects exhibited slowing only when ITI was short (short ITI: $\Delta RT = 54 \pm 13.3$ ms; long ITI: $\Delta RT = 16 \pm 19.9$ ms). (E) PES persists across multiple trials only when ITI is short. ΔRT indicates the difference between post-correct (pre-error) RTs and the mean RT one, two, or three trials after the error. When ITI is short, PES persists for two or three trials. When ITI is long, PES is eliminated after a single trial. The figure aggregates data across four subjects that showed PES in the ITI manipulation experiment. A similar ITI dependence exists for post-error speeding. In a fifth subject that showed speeding, speeding declined faster across trials for longer ITI (Data not shown). Asterisks indicate results of linear regression analysis (Equations 2 and S1; *** for $p < 0.001$, ** for $p < 0.01$, * for $p < 0.1$).

Table S1, related to Figure 1. Effects of motion strength and previous trial feedback on individual subject response times (Equation 2) and proportion correct (Equation 1).

	β_1 (motion strength, %Coh)		β_2 (previous-trial feedback)		β_3 (interaction)	
Response time (ms)						
Humans						
S1	-1249.84 ±	21.46 ***	131.61 ±	27.72 ***	-235.74 ±	202.09
S2	-1814.75 ±	35.05 ***	150.91 ±	44.11 **	-493.03 ±	325.45
S3	-3115.35 ±	38.28 ***	138.08 ±	50.62 **	-634.17 ±	383.65
S4	-1379.71 ±	34.54 ***	197.38 ±	46.67 ***	-719.66 ±	343.49 *
S5	-1511.92 ±	27.48 ***	71.34 ±	36.34	-139.27 ±	253.61
S6	-2555.39 ±	55.90 ***	379.77 ±	72.94 ***	-1165.08 ±	524.01 *
S7	-2620.97 ±	37.29 ***	-165.92 ±	49.33 **	652.37 ±	360.61
Monkeys						
O	-197.88 ±	19.25 ***	68.53 ±	9.78 ***	-60.33 ±	26.79 *
N	-963.04 ±	72.48 ***	114.15 ±	25.76 ***	-174.68 ±	95.10
D	-303.03 ±	21.99 ***	-97.61 ±	8.48 ***	195.79 ±	30.60 ***
B	-580.60 ±	43.22 ***	-1.26 ±	16.28	-48.13 ±	58.39
	β_1 (motion strength, %Coh)		β_2 (previous-trial feedback)			
Proportion correct						
Humans						
S1	14.30 ±	1.72 ***	-0.25 ±	2.44		
S2	16.62 ±	2.00 ***	-0.61 ±	2.74		
S3	38.92 ±	5.69 ***	-14.84 ±	6.75 *		
S4	25.32 ±	3.42 ***	-0.46 ±	4.99		
S5	20.20 ±	2.38 ***	-2.95 ±	3.12		
S6	19.37 ±	2.73 ***	-4.32 ±	3.47		
S7	36.35 ±	4.70 ***	-4.12 ±	6.36		
Monkeys						
O	11.96 ±	0.73 ***	-0.86 ±	1.00		
N	26.05 ±	4.30 ***	-5.18 ±	5.01		
D	10.05 ±	0.61 ***	-1.63 ±	0.78 *		
B	20.59 ±	2.64 ***	-1.76 ±	3.32		

Each row shows the coefficients of Equation 1 (proportion correct) or Equation 2 (response time). *** denotes $p < 0.001$, ** denotes $p < 0.01$, * denotes $p < 0.05$.

Table S2, related to Figures 3 and 4. Best-fitting drift-diffusion model parameter values to individual subject data.

	S1	S2	S3	S4	S5	S6	S7	Monkey O	Monkey N	Monkey D	Monkey B
\bar{b}											
post-correct	22.1 ± 1.05	30.4 ± 0.82	32.0 ± 0.79	24.1 ± 0.86	23.1 ± 0.78	27.8 ± 0.95	29.0 ± 0.88	15.4 ± 0.86	26.0 ± 1.02	14.1 ± 0.37	20.0 ± 0.59
post-error	24.3 ± 0.86	32.2 ± 0.69	34.6 ± 0.66	26.8 ± 0.98	24.1 ± 0.93	32.2 ± 0.84	24.1 ± 1.88	16.5 ± 0.96	26.5 ± 0.70	13.7 ± 0.27	21.3 ± 0.63
k											
post-correct	0.47 ± 0.11	0.29 ± 0.03	0.52 ± 0.04	0.46 ± 0.05	0.49 ± 0.05	0.48 ± 0.04	0.65 ± 0.04	0.30 ± 0.03	0.56 ± 0.08	0.43 ± 0.04	0.44 ± 0.06
post-error	0.40 ± 0.09	0.26 ± 0.02	0.53 ± 0.04	0.42 ± 0.05	0.45 ± 0.06	0.36 ± 0.04	0.56 ± 0.06	0.27 ± 0.02	0.48 ± 0.06	0.31 ± 0.02	0.62 ± 0.05
t_0											
post-correct	337.6 ± 66.32	214.9 ± 32.19	91.4 ± 22.97	345.0 ± 35.68	295.5 ± 29.65	231.3 ± 17.59	283.5 ± 9.40	360.7 ± 15.55	241.8 ± 48.68	353.6 ± 7.82	324.9 ± 18.31
post-error	335.8 ± 62.73	192.3 ± 33.08	85.2 ± 14.17	338.6 ± 34.44	320.8 ± 35.12	165.9 ± 15.45	273.7 ± 11.70	372.7 ± 21.35	262.9 ± 20.21	310.7 ± 7.94	333.8 ± 11.72
$\tau_{1/2}$											
post-correct	168.1 ± 56.13	459.5 ± 105.88	545.3 ± 105.78	1977.3 ± 795.66	1018.6 ± 14064.67	0.9 ± 1643.75	16978.9 ± 6074.52	221.4 ± 53.88	207.0 ± 39.58	809.4 ± 21186.91	63.2 ± 5.58
post-error	237.7 ± 70.77	591.8 ± 110.97	703.5 ± 148.68	4140.0 ± 1658.16	189.5 ± 4907.10	6921.7 ± 3501.17	194.6 ± 31680.53	296.6 ± 50.97	231.1 ± 32.27	0.1 ± 1968.08	65.6 ± 4.80
u_{∞}											
post-correct	88.3 ± 17.52	88.0 ± 12.78	88.2 ± 8.52	32.1 ± 5.11	6.8 ± 6.80	10.7 ± 3.19	16.5 ± 5.92	32.8 ± 21.27	88.7 ± 20.02	18.2 ± 3.73	126.0 ± 0.17
post-error	88.3 ± 17.56	88.9 ± 13.70	88.2 ± 8.83	32.1 ± 6.41	2.3 ± 5.77	37.7 ± 4.44	7.8 ± 5.79	32.8 ± 22.48	88.7 ± 19.84	4.9 ± 2.51	126.0 ± 0.08
b											
all trials	90.2 ± 17.78	90.8 ± 13.40	90.1 ± 8.70	32.7 ± 4.91	25.8 ± 1.88	38.5 ± 2.45	29.9 ± 0.91	33.5 ± 21.56	90.6 ± 20.25	18.5 ± 0.54	128.6 ± 28.44
$t_{0\text{ bias}}$											
all trials	110.6 41.53	-190.4 ± 55.39	47.2 ± 13.067	-85.4 ± 55.1997	-11.8 ± 50.52	54.8 ± 21.34	-32.0 ± 15.03	128.1 ± 23.92	-2.2 ± 19.321	-4.1 ± 2.257	-18.2 ± 24.92
z											
all trials	0.0 ± 0.03	0.2 ± 0.04	-0.1 ± 0.03	0.2 ± 0.08	0.0 ± 0.12	0.0 ± 0.05	0.1 ± 0.04	0.2 ± 0.03	0.1 ± 0.02	-0.1 ± 0.07	0.0 ± 0.03
S_{t0}											
all trials	79.9 ± 38.34	2.2 ± 2.42	1.0 ± 0.34	46.3 ± 16.90	23.7 ± 11.88	2.9 ± 1.55	3.4 ± 1.41	69.1 ± 6.03	99.4 ± 14.33	50.1 ± 2.93	76.8 ± 6.63

Each row shows best-fitting parameter values (±SE) for drift-diffusion model fits to individual subject data (see Experimental Procedures). Standard error was computed using a nonparametric bootstrap. \bar{b} indicates the average bound over the middle 95% of decision times (see Figure 4). Starting point (z) is reported as a proportion of b .

SUPPLEMENTAL EXPERIMENTAL PROCEDURES

Additional behavioral task details

Both human and monkey subjects were extensively trained so that all subjects, regardless of species, achieved high accuracy and stable speed (see Results). For the final data collection, we emphasize both speed and accuracy and subjects exhibited intermediate RTs and psychophysical thresholds comparable to previous reports (Hanks et al., 2014; Palmer et al., 2005). Subjects could exhibit higher accuracies when instructed to do so, indicating that the lack of post-error change of accuracy was not due to a ceiling effect. To further establish this point, we performed a control analysis by dividing experimental sessions into two groups based on psychophysical thresholds (median split). Our results were similar for both high- and low-threshold sessions, which again rules out ceiling effects on post-error accuracy.

To achieve the high trial counts required for the analyses, subjects contributed several sessions of data. In each session, humans performed several blocks of 100-200 trials (1,959 – 2,846 total trials per subject, mean = 2,427 trials, total = 16,989 trials). Monkeys performed the task for liquid reinforcement until satiated in each session (total trial count, monkey N: 3,534; monkey O: 21,179; monkey B: 2,615; monkey D: 15,290). Stimulus presentation was controlled with Psychophysics Toolbox (Brainard, 1997) and Matlab. Eye movements were monitored using a high-speed infrared camera (Eyelink, SR-Research, Ontario) for humans and monkey O. A scleral search coil (Fuchs and Robinson, 1966; Judge et al., 1980) was used to monitor eye movements of monkeys N, B, and D. Gaze positions were recorded at 1kHz.

The direction and motion strength was randomly chosen on each trial. For monkeys, motion strength could be one of the following values: 0%, 3.2%, 6.4%, 12.8%, 25.6%, and 51.2%. Because our human subjects were better at the task than monkeys, we used a slightly more difficult stimulus set: 0%, 1.6%, 3.2%, 6.4%, 12.8%, and 25.6%. This ensured a comparable range of difficulty for all subjects.

Following the response, the humans received distinct auditory tones indicating whether the response was correct or an error. Monkey B and N received a tone and liquid reward after correct responses. Monkey O and D received immediate feedback (tone + liquid for corrects, different tone for errors) on some trials. For the other trials, the initial choice was followed by the presentation of two distinct gambling targets. Shifting gaze to one gambling target would result in auditory feedback and a large reward following correct direction choices or no reward for incorrect direction choices (i.e., high-risk target). Shifting gaze to the other gambling target would result in a small reward regardless of the initial choice (i.e., low-risk target). Our focus here is to understand the effect of errors on the primary motion discrimination task, and therefore we only analyzed behavior following trials in which explicit feedback was given (i.e., non-gambling trials and high-risk choices). For all subjects, trials with 0% coherent motion were rewarded randomly. We defined these trials as correct or error based on the feedback that was received.

Definition of post-correct trials

We used only post-correct trials that are also pre-error trials to guard against the possibility that fluctuations in behavioral state could create confounds across PES conditions (Dutilh et al., 2012; Nelson et al., 2010). This additional criterion is not critical for our results, but it ensures that post-error trials are matched to the post-correct trials that occurred around the same time in the session. This has been demonstrated to protect against the influence of performance fluctuations (Dutilh et al., 2012). We observed very similar results even when all post-correct trials were considered, suggesting that performance fluctuations were not a major factor in our data. In addition, we repeated our behavioral analyses using only post-correct trials that were followed by another correct response to guard against the possibility that our results were influenced by systematic effects that precede errors (Allain et al., 2004; Eichele et al., 2008; Ridderinkhof et al., 2003). Again, the results were very similar suggesting that selection of post-correct trials that were pre-error was uncritical for our results.

Inter-trial interval manipulations

If PES is caused by adaptive increases in decision bound to compensate for reduced sensitivity following negative feedback, then PES should be sensitive to the delay between negative feedback and the next decision (Danielmeier and Ullsperger, 2011; Dudschig and Jentzsch, 2009; Jentzsch and Dudschig, 2009). Specifically, extending the inter-trial interval (ITI) would dissipate the effects of negative feedback and provide time for restoring normal sensory sensitivity. To test this prediction, we conducted a follow up experiment in which we varied ITI across blocks for five human subjects. Each subject experienced two ITI conditions; a short ITI block and a long ITI block. Each ITI condition was completed across multiple consecutive sessions before switching to the new condition. For two subjects, the short ITI was 0.8s and the long ITI was 3.2s. For three subjects, the short ITI was 0.35s and the long ITI was 2.4s. The ordering of the conditions was counterbalanced across subjects.

We used Equations 1 and 2 to test for significant changes in RT and proportion correct following errors in both ITI conditions. To test the significance of the effect of ITI on the magnitude of PES, we used the following regression equation:

$$\Delta RT = \beta_0 + \beta_1 I \quad (\text{Equation S1})$$

where I is an indicator variable set to 1 for the short ITI condition and 0 otherwise. ΔRT is the difference between post-error RT on individual trials and the post-correct/pre-error RT that preceded it (total number of trials for short ITI: $n = 3,932$; long ITI: $n = 3,140$ trials).

Model fitting procedure

For each fit, we calculated the probability of the accumulated evidence crossing the upper or lower bound at each decision time using a numerical solution for the Fokker-Planck equation (Kiani et al., 2014; Kiani and Shadlen, 2009). We fit the full RT distributions by maximizing the likelihood of RTs on individual trials irrespective of choice. A nonparametric bootstrapping method (100 iterations) was employed to estimate the standard error of parameter estimates. Because our fits ignored choice, we could then use the fitted parameters to generate predictions for the choice probabilities associated with each condition.

To probe the cause of PES, we separately fit the observed data from post-correct and post-error trials. We adopted a two-stage fitting approach in order to reduce computational demands and ensure that our fits converged to a global minimum. First, we optimized all eight parameter values to fit post-error RTs. Next, we refit key parameters of interest to the post-correct RTs while fixing parameters not likely to explain PES. We fixed both bias parameters (z , and t_{0bias}) across post-correct and post-error conditions because we found that RT biases did not change following errors. In addition, we fixed s_{t0} across post-correct and post-error conditions because this parameter reflects variability inherent in the sensory and motor systems. Overall, changes in five model parameters were explored as the cause for PES (b , k , t_0 , u_{∞} , $\tau_{1/2}$). Our results did not depend critically on whether we fit post-error or post-correct data first, or whether we shared b , s_{t0} , z , and t_0 across the post-correct and post-error fits.

We used weighted least-squares regression to assess the significance of model parameter changes across subjects. The regression equation was defined by the following function:

$$\Delta P = \beta + \varepsilon \quad (\text{Equation S2})$$

where ΔP is a vector of differences in parameter values fit to post-error and post-correct data for each subject. β was found by minimizing the function $S = \sum_{i=1}^n w_i (\Delta P_i - \beta)^2$, where w_i are individual subject's weights (n subjects) given by the inverse of the variance of model parameter differences estimated from bootstrapping. The null hypothesis is no change following errors ($H_0: \beta = 0$). To test overall changes in bound height in the presence of urgency, we averaged $B(t)$ (Equation 4) between the 2.5th to 97.5th percentile of the decision time distribution. The resulting mean, \bar{B} , corresponds to the average bound

height when most choices occurred. Equation S2 was then used to test for overall changes in the mean bound height across post-correct and post-error trials.

If unexpected negative feedback drives reductions in sensitivity after errors one would expect that the magnitude of post-error sensitivity reductions depends on error-trial motion strength. We tested this hypothesis by evaluating a modified drift-diffusion model that allowed rescaling of post-error drift rate based on error-trial motion strength. In this model, the drift rate was $\mu = kC_{cur}$ for post-correct trials, and $\mu = (k + \alpha C_{prev}) \times C_{cur}$ for post-error trials, where C_{cur} and C_{prev} are the motion strength on the current and previous trials, respectively. Negative α indicates that sensitivity is reduced more when negative feedback is less expected.

Changes in bound height with PES could be due to either a static offset in overall bound, b , or a dynamic change in urgency (the rate of the bound collapse parameterized by u_{∞} and $\tau_{1/2}$). We used nested model comparisons to test whether static or dynamic bound changes were necessary to explain PES. We used a likelihood-ratio test to determine the necessity of each mechanism. For comparisons of non-nested models, we used a parametric bootstrapping procedure. As explained in the Results, we found that changes in urgency (the rate of bound collapse) were necessary to explain PES, and so we focus primarily on this model.

Neural recording details

Neuronal activity was recorded from the lateral intraparietal area (LIP) of four rhesus monkeys while they performed the motion discrimination task (50 neurons from Monkey O; 64 neurons from Monkey N; 37 neurons from Monkey D; 52 neurons from Monkey B). LIP was located based on stereotaxic coordinates and structural magnetic resonance imaging (MRI) scans. Neurons were selected for recording if they exhibited spatially selective persistent activity during the delay period of memory-guided and visually-guided delayed saccade tasks (Gnadt and Andersen, 1988; Hikosaka and Wurtz, 1983). These tasks require the animal to maintain fixation following target presentation throughout a delay period (~1000ms) until the offset of the fixation point cues the animal to initiate a saccade to the target (visually-guided saccades) or remembered target location (memory-guided saccades). The target location was varied across trials to identify the spatial location that elicited a maximal response from the neuron. This location was designated the response field (RF).

Estimation of urgency signal

We used two different analyses to test whether PES is accompanied by evidence-independent changes in the temporal dynamics of LIP responses. First, following Churchland et al. (2008), we averaged responses across all post-correct trials regardless of motion direction, motion strength, choice, and outcome. In this average, changes in activity related to incoming perceptual evidence cancel out, and the remaining modulation of activity is due to a nonspecific signal that accelerates commitment to a choice. To combine responses across neurons, we normalized the firing rate of each neuron to the average firing rate in a 100ms window before dots onset. The urgency signal was independently calculated for the post-correct and post-error trials. Responses were truncated 100 ms prior to saccade initiation, to ensure that the approximated urgency was not influenced by pre-saccadic response bursts. Similar results were obtained when we quantified urgency only based on low coherence stimuli ($\leq 6.4\%$), which had little net motion information.

The post-correct and post-error urgency signals were parameterized with a hyperbolic function fit to the epoch from the response dip until 450 ms after dots onset (Figure 6A):

$$R(t) = \beta_0 + u_{\infty} \frac{t}{t + \tau_{1/2}} \quad (\text{Equation S3})$$

where $R(t)$ indicates the urgency at time t , β_0 is the baseline, and u_{∞} and $\tau_{1/2}$ determine asymptotic level and half time of the urgency. We computed the best fitting parameter values for individual neurons and

tested for significant changes between post-correct and post-error trials using Wilcoxon signed-rank tests. Similar results were obtained when we fit the same function to the population average responses (Figure 6A).

In the second analysis, following Hanks et al. (2011), we estimated the urgency signal by computing the buildup rate on a trial-by-trial basis. We divided the trial using a running window of 150ms with 25ms time steps from 100 to 400 ms after dots onset for post-correct and post-error trials. Bins that included within 100 ms before the time of saccade were excluded. We computed the buildup rate in each bin by convolving the spike train with a Gaussian kernel (SD = 25ms) and then by applying a linear regression to this smoothed rate function. The average buildup rate across all trials, irrespective of choice, provides a piecewise linear approximation to the first derivative of the urgency signal. The time integral of this derivative provides an estimate of the dynamic component of the urgency signal itself (Figure 6B). We used a parametric bootstrap (1000 samples) to estimate standard error (Hanks et al., 2011) and assess significant differences in the neurally-derived urgency for post-correct and post-error trials on a bin-by-bin basis.

VarCE analysis

The analysis was performed using a 60 ms window that was moved in steps of 10 ms. We estimated the point process variance of each neuron at each time by multiplying the spike count with an upper bound estimate of the neuron’s Fano Factor (ϕ): the largest possible value that ensured a positive VarCE at all times. For each neuron we computed the spike count in each time step and subtracted the mean spike count across all trials with a similar condition (e.g., motion strength, direction, etc). To enhance the analysis power we combined these conditional residual spike counts across neurons and computed VarCE in each time bin as the total variance of the residuals from all conditions and neurons, $Var[z_U]$, minus the weighted average of the point process variance for each condition:

$$VarCE = Var[z_U] - \sum_{j=1}^K \sum_{i=1}^M \frac{n_{ij}}{n_U} \phi_j \bar{N}_{ij} \quad (\text{Equation S4})$$

where n_U is the total number of trials across M conditions and K neurons, n_{ij} is the number of trials, and \bar{N}_{ij} is the mean spike count for the i^{th} condition and j^{th} neuron. A common ϕ_j was used across all conditions for a neuron. It was computed as the minimum value of the Fano factor across the trial, and following Equation S5 the Fano Factor for each neuron was calculated as:

$$FF = \frac{Var[z_U]}{\sum_{i=1}^M \frac{n_i}{n_U} \bar{N}_i} \quad (\text{Equation S5})$$

where z_U and n_U refer to all recorded trials for the neuron. We verified that ϕ_j was stable across post-correct and post-error conditions, indicating that the point-process noise does not change. Standard error of VarCE was estimated using a bootstrap procedure (250 iterations, random sampling of trials with replacement). The definition of conditions depended on the epoch of interest. For early decision formation (motion aligned VarCE, Figure 7C, left panel) they were defined based on motion direction and strength. For the perisaccadic epoch (saccade aligned VarCE, Figure 7C, right panel), they were defined based on motion direction, motion strength, and choice.

We used a nonparametric bootstrap to test for significant differences in VarCE and Fano factor for post-correct and post-error trials. For each time bin, we computed the difference in neural variance between post-error and post-correct trials, and generated 95% confidence intervals. Our results were identical if we computed VarCE and Fano factor for each neuron individually then tested for significant differences between post-correct and post-error trials across neurons.

SUPPLEMENTAL REFERENCES

- Allain, S., Carbonnell, L., Falkenstein, M., Burle, B., and Vidal, F. (2004). The modulation of the Ne-like wave on correct responses foreshadows errors. *Neurosci Lett* 372, 161-166.
- Brainard, D.H. (1997). The Psychophysics Toolbox. *Spat Vis* 10, 433-436.
- Dudschig, C., and Jentzsch, I. (2009). Speeding before and slowing after errors: is it all just strategy? *Brain Res* 1296, 56-62.
- Dutilh, G., van Ravenzwaaij, D., Nieuwenhuis, S., van der Maas, H.L., Forstmann, B.U., and Wagenmakers, E.-J. (2012). How to measure post-error slowing: a confound and a simple solution. *Journal of Mathematical Psychology* 56, 208-216.
- Eichele, T., Debener, S., Calhoun, V.D., Specht, K., Engel, A.K., Hugdahl, K., von Cramon, D.Y., and Ullsperger, M. (2008). Prediction of human errors by maladaptive changes in event-related brain networks. *Proc Natl Acad Sci U S A* 105, 6173-6178.
- Fuchs, A., and Robinson, D. (1966). A method for measuring horizontal and vertical eye movement chronically in the monkey. *Journal of Applied Physiology* 21, 1068-1070.
- Gnadt, J.W., and Andersen, R.A. (1988). Memory related motor planning activity in posterior parietal cortex of macaque. *Exp Brain Res* 70, 216-220.
- Hikosaka, O., and Wurtz, R.H. (1983). Visual and oculomotor functions of monkey substantia nigra pars reticulata. III. Memory-contingent visual and saccade responses. *J Neurophysiol* 49, 1268-1284.
- Jentzsch, I., and Dudschig, C. (2009). Why do we slow down after an error? Mechanisms underlying the effects of posterror slowing. *Q J Exp Psychol (Hove)* 62, 209-218.
- Judge, S.J., Richmond, B.J., and Chu, F.C. (1980). Implantation of magnetic search coils for measurement of eye position: an improved method. *Vision Research* 20, 535-538.
- Kiani, R., and Shadlen, M.N. (2009). Representation of confidence associated with a decision by neurons in the parietal cortex. *Science* 324, 759-764.
- Nelson, M.J., Boucher, L., Logan, G.D., Palmeri, T.J., and Schall, J.D. (2010). Nonindependent and nonstationary response times in stopping and stepping saccade tasks. *Atten Percept Psychophys* 72, 1913-1929.
- Ridderinkhof, K.R., Nieuwenhuis, S., and Bashore, T.R. (2003). Errors are foreshadowed in brain potentials associated with action monitoring in cingulate cortex in humans. *Neurosci Lett* 348, 1-4.