Neuron, Volume *83*

Supplemental Information

Effects of Cortical Microstimulation on Confidence in a Perceptual Decision

Christopher R. Fetsch, Roozbeh Kiani, William T. Newsome, and Michael N. Shadlen

Index

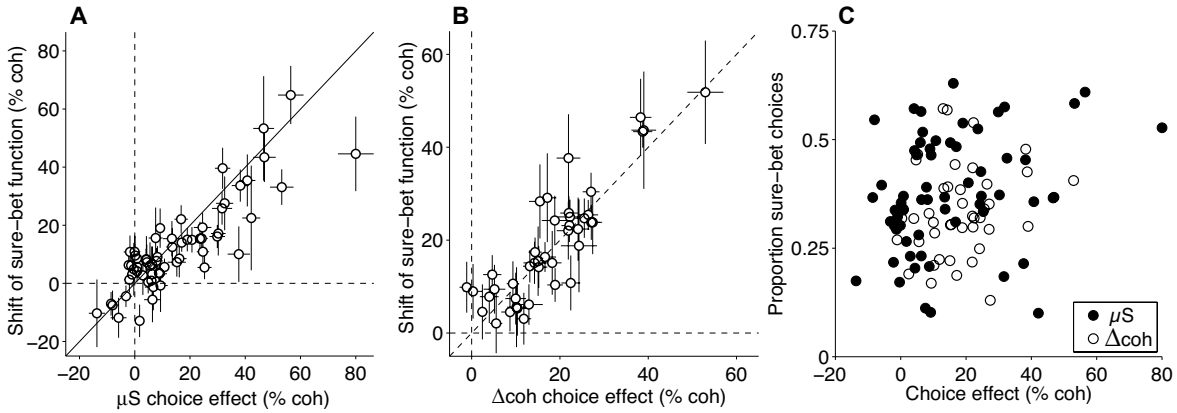

Figure S1 [related to Figure 2B]. Comparison of microstimulation effects on choice and PDW using descriptive fits (logistic regression and Gaussian, respectively; see Experimental Procedures). (**A**) For each stimulation site (n = 63), the effect on choice is expressed as a horizontal separation of the sigmoid functions from logistic regression (β_1/β_2 in Equation 1, Experimental Procedures). The effect on PDW is the shift of the fitted Gaussian functions (β_2 in Equation S2). Error bars are SE of the estimates. Pearson's $r > 0.8$ for each brain area and each monkey separately. (**B**) Same format as A, but for Δcoh control sessions (n = 42 experiments). (**C**) Variation in the proportion of sure bet choices across experiments is not explained by the size of μS and Δcoh effects on choice (Pearson $r < 0.02$, $P > 0.9$ for both). The graph also demonstrates that the levels of confidence were comparable across μS and Δcoh experiments ($P = 0.67$, 2-dimensional Kolmogorov-Smirnov test).

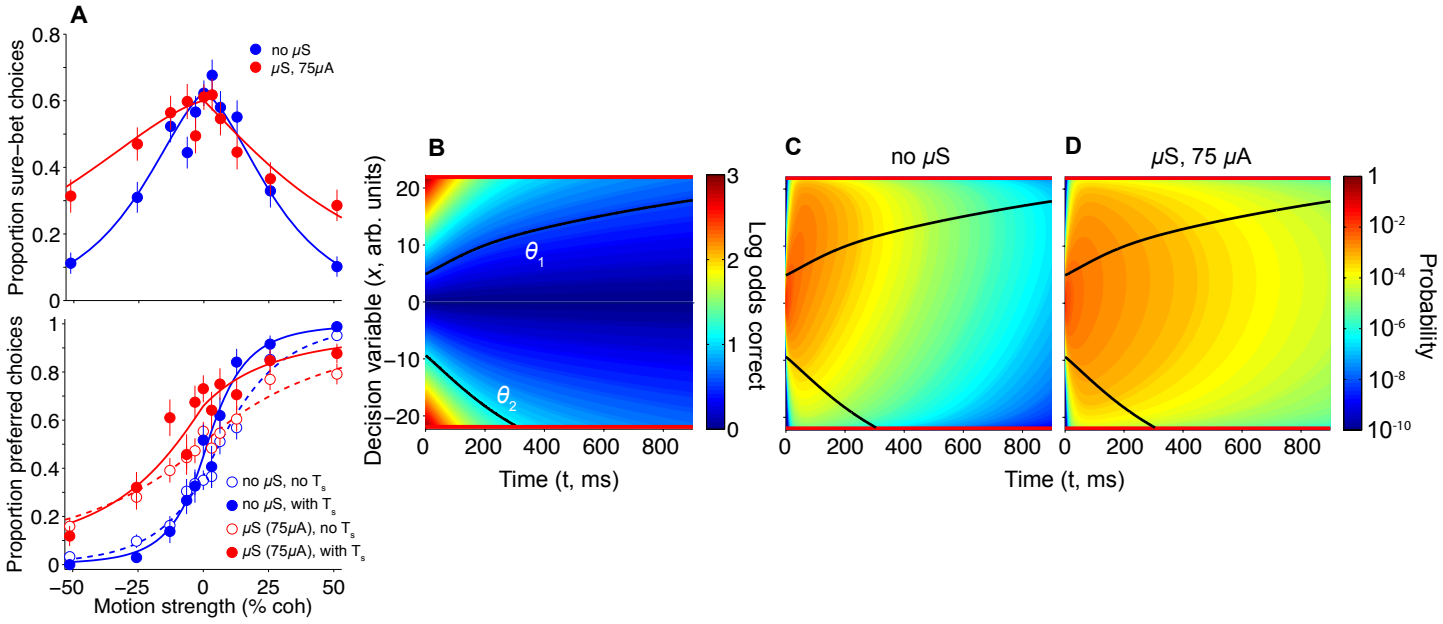

Figure S2 [related to Figure 3 and Suppl. Experimental Procedures]. Extended model used to explain high-current stimulation experiments. (**A**) Data and fits (n = 4,483 trials), reproduced from main text Figure 3. (**B-D**) Depiction of the extended model (see p. 13). Heat map in **B** displays the certainty that ought to accompany a decision based on the sign of the accumulated evidence *x* at decision time t. Units are the natural logarithm of the odds of a correct choice. The mapping is derived identically to the one in main text Figure 4B, independent of valuation, bias, or μS. It depends only on the possible motion strengths and the parameter α. The criteria for a sure-bet choice, θ_1 and θ_2, were allowed to differ for preferred and null choices (upper and lower halves of the map). (**C,D**) The probability density of the decision variable across time is shown for 51.2% coherence motion in the preferred-direction, separated for no-μS (**C**) and high-current μS (**D**) conditions. Without μS, the density is shifted strongly upward, generating a high proportion of preferred-direction choices and relatively few sure-bet choices. The effect of high-current μS is mainly to reduce sensitivity by activating opposing direction columns, implemented in the model by altering the relationship between coherence and drift rate (δ_α in Equation S14; see Table S1). Thus, even for the highest motion strength (51.2% coh), the model produces a decrease in accuracy and an increase in the proportion of sure-bet choices, as observed in the data (panel A, top). In addition, the asymmetric confidence region (difference between θ_1 and θ_2, panel B) generates a peculiar bias similar to the one observed in the data: more preferred direction choices on T_s-waived relative to no-T_s trials (panel A bottom, solid curves vs. dashed counterparts).

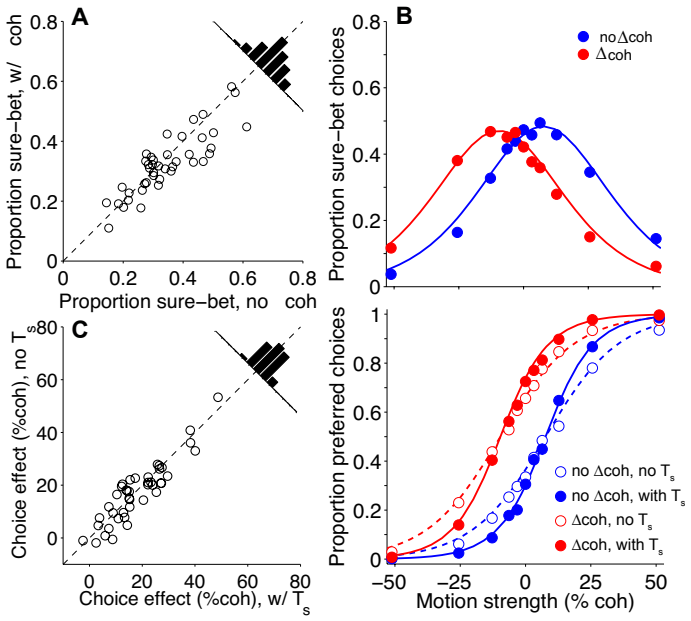

Figure S3 [related to Results: Controls and Figure 2B top, inset]. Similar pattern of PDW and choice behavior with a coherence offset added to the motion instead of μS. (A–C) Same format as main text Figure 2, but for control ('Δcoh') sessions without electrical μS. Top panel of B is the same as inset in Figure 2B, top. The coherence offset (fixed within sessions, varied between 5-40% across sessions) was added to a random half of trials in an arbitrary preferred direction. Trial structure and reward contingencies were identical to μS sessions.

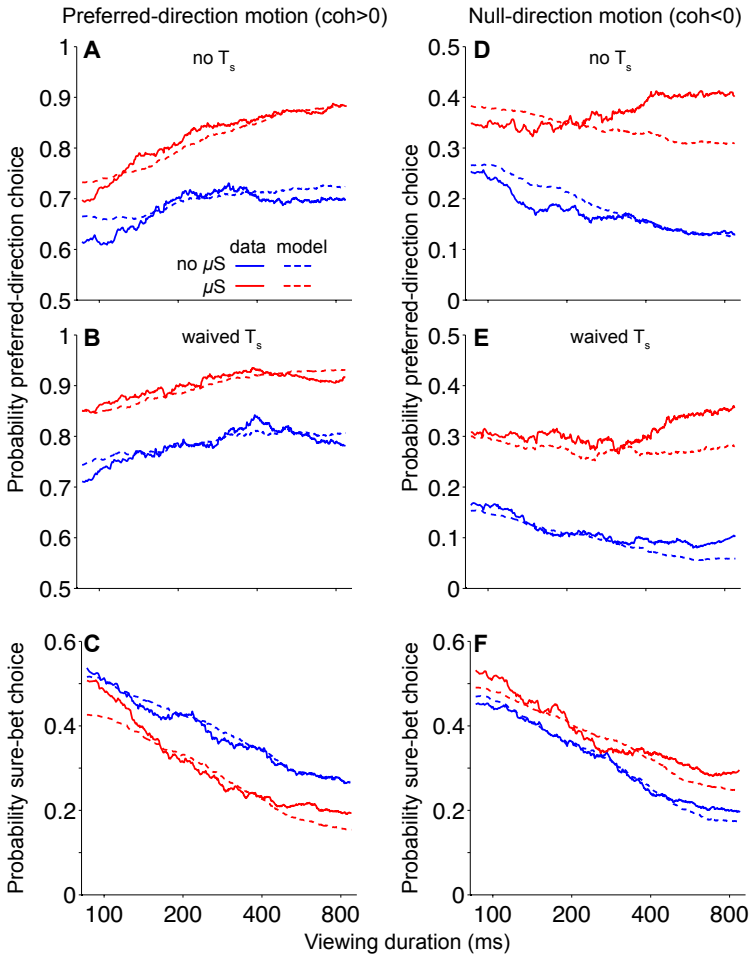

Figure S4 [related to Experimental Procedures and Figure 4]. Model fits to direction choices and sure-bet choices in the main μS experiments, separated by motion direction. (**A-C**) For preferred-direction motion (coh>0), μS increases the probability of a preferred choice (**A** & **B**, red vs. blue) and decreases the probability of a sure-bet choice (**C**). (**D-F**) μS also increases preferred-direction choices when motion is in the null direction (coh<0; **D** & **E**), and increases the probability of a sure-bet choice as well (**F**). Solid traces are running means (proportions) of the data sorted by viewing duration; dashed traces are fits to the simpler bounded accumulation model described in the main text and in detail below. Implementing μS in the model as an equivalent change in motion strength reproduces the basic pattern of effects as a function of viewing duration. One exception is that the model underestimates the probability of a preferred direction choice for coh<0 and long durations (red curves in **D** & **E**). It appears that longer microstimulation epochs have a greater ability to counteract opposing visual motion than predicted by a static change in motion strength.

Table S1 [related to Figure 3 and Table 1]. Maximum-likelihood estimates of extended model parameters (± s.e.). Note that the extended model recovers the main features of the simpler model used to fit the low current μS experiments (repeated here from Table 1 for ease of comparison).

	High-current μS, extended model	Low-current μS, extended model	Low-current μS, basic model
α	0.345 ± 0.017	0.327 ± 0.010	0.294 ± 0.001
B	21.9 ± 4.95	30.7 ± 9.10	31.6 ± 1.30
γ	-0.021 ± 0.001	-0.016 ± 0.001	-0.013 ± 0.000
θ_1	0.561 ± 0.400	0.551 ± 0.013	0.609 ± 0.004
θ_2	1.118 ± 0.075	0.623 ± 0.017	N/A
δ_C	0.142 ± 0.006	0.129 ± 0.003	0.112 ± 0.001
δ_{σ^2}	0.020 ± 0.049	0.156 ± 0.027	0.237 ± 0.015
δ_α	-0.190 ± 0.010	-0.040 ± 0.001	N/A
β	1.48 ± 0.54	0.91 ± 0.32	N/A

SUPPLEMENTAL EXPERIMENTAL PROCEDURES

Neurophysiological methods and selection of stimulation sites

Physiological criteria for identifying MT and MST included (a) audible transitions between white and gray matter, (b) vigorous responses to patches of moving dots, and (c) receptive field (RF) size vs. eccentricity. Stimulation sites classified as MT had RF diameters roughly equal to the eccentricity of the RF center and which rarely included the fovea, whereas MST sites had larger RFs which frequently included the fovea and/or part of the ipsilateral hemifield. When the angle of penetration allowed both MT and MST to be accessed in the same electrode track, we confirmed the depths of both areas as well as the white matter below the superior temporal sulcus, either before or after the behavioral session as time permitted. One monkey was euthanized and histological sections were processed for immunohistochemistry using the SMI32 antibody (neurofilament protein), which produces a characteristic staining pattern in area MT (Hof and Morrison, 1995). Examination of these sections — which contained visible guide tube tracks and an injection site placed near the microstimulation sites (for a separate study) — and comparison to brain atlases confirmed that all sites classified as MT in this animal (N=21) were indeed located in the posterior-medial portion of MT.

We isolated multiunit (MU) activity using a dual voltage-time window discriminator (BAK Electronics, Sanford, FL), with voltage and slope thresholds configured to include events slightly above the background 'hash' (typical pre-stimulus baseline event rate = 30-100 events/s). Once the electrode was positioned near a site with strong visual responses to flashing random-dot patches, we measured MU tuning for motion direction, speed, patch size, and position while the monkey fixated a central target. Parametric functions were fit to the MU event rates (circular Gaussian and log-normal for direction and speed tuning, respectively), and the preferred direction and speed were taken as the peaks of their respective fitted functions. Preferred size (approximate RF diameter) was taken as the patch size that elicited the strongest response, or in

the case of saturating/sigmoidal size tuning, as the smallest value in the saturated portion of the tuning function.

Tuning measurements were repeated 2-4 times at intervals of 80-120 μm along the electrode track. A site was considered acceptable for microstimulation if (a) tuning/RF parameters remained relatively stable (e.g., Δ pref. dir. < 50°) for at least 200 μm, and (b) if direction selectivity was sufficiently strong, with at least 2 standard deviations separating preferred- and null-direction MU responses.

Addendum to behavioral data analysis

We fit the probability of sure-bet choices as function of signed coherence with a Gaussian function of the form:

$$P_{sb}(C) = Ae^{-(C-\mu)^2/(2\sigma^2)} \qquad \text{(Equation S1)}$$

The three free parameters (A, μ, σ) and their standard errors were estimated using the same methods as Equation 1 (main text). To evaluate the significance of changes in each parameter with μS, we fit an expanded function with three additional free parameters β_i:

$$P_{sb}(C) = (A + \beta_1 I_E) \cdot e^{-(C-\mu+\beta_2 I_E)^2/(2\cdot(\sigma+\beta_3 I_E)^2)} \qquad \text{(Equation S2)}$$

As in Equation 1, I_E is the μS indicator variable (0 or 1). Having found that the primary effect of μS was a change in the mean (i.e., a horizontal shift; $\beta_2 > 0$), we quantified this shift (e.g., for comparison with β_1/β_2 in Equation 1; see Figures S1A and S1B) using a simplified version of Equation S2 with β_1 and β_3 set to zero. Rank correlations reported in the main text (Results: Controls) refer to the following comparisons: σ in Equation S1 versus β_2 in Equation 1, and β_3 in Equation S2 versus β_3 in Equation 2.

To compare the frequencies of T_s choices with vs. without μS (or Δcoh), we used a paired t-test on logit-transformed proportions. The raw proportions are shown in Figures 2A and S3A, but statistical tests were performed after correcting for a potential artifact arising from (*i*) the nonuniform sampling of coherences (e.g., points concentrated around 0% coh) and (*ii*) the fact

that the shifts were not fully counteracted by the compensatory bias. For each session, we effectively slid the coherence axis so that 0% was equidistant from the means of the fitted Gaussians (Equation S1) for μS and no-μS (or Δcoh and no-Δcoh) conditions. The Gaussians provided interpolated proportions at the log-spaced coherences (e.g., 0%, ±3.2%, ..., ±51.2%) on the shifted abscissa. Averaging these proportions yielded a corrected estimate of sure-bet frequency, free from sampling confounds. Results of paired t-tests on the corrected values suggested no change in the overall sure-bet frequency for Δcoh vs. no-Δcoh (P=0.88; Figure S3A) and a decreasing, nonsignificant trend for μS vs. no-μS (P = 0.18). To better quantify this trend, we examined the change in the height parameter, A, of the fitted Gaussians across sessions. This comparison showed a significant decrease in the height of the sure-bet function for μS (paired t-test, P < 0.004), but, as expected, no change for Δcoh (P = 0.47). Incidentally, the decrease in A did not entail a reduction in (corrected) sure-bet frequency because it was compensated by a small increase in width (σ). All of these subtle details are well explained by the bounded accumulation model, in particular the addition of noise on μS trials (δ_{σ^2} ; see Table 1 and above text).

Detailed description of bounded accumulation model and tiered fitting procedure

As described briefly in Experimental Procedures, we modeled the results using a simple one-dimensional bounded-accumulation (or drift-diffusion) model. Although some task variants (e.g., reaction time studies) require more complex models, such as nonstationary bounds and races between competing processes, here we opted for the simplest model framework capable of explaining our results.

The accumulated evidence (or decision variable, x) undergoes deterministic drift plus diffusion, conforming to the discrete stochastic difference equation

$$\Delta x = R\Delta t + N\left\{0, \sqrt{\sigma^2 \Delta t}\right\}, \quad x(0) = 0 \qquad \text{(Equation S3)}$$

where R is the drift rate and $N\{\cdots\}$ is a Normal (Gaussian) distribution with mean 0 and standard deviation $\sigma\sqrt{\Delta t}$. In the absence of bias and microstimulation, *R* is proportional to signed motion coherence ($R = \alpha C$; $\alpha > 0$), where positive values of *C* indicate motion in the preferred direction, and the variance $\sigma^2 = 1$. The process terminates when the stimulus is turned off ($t_d = \tau$) or when *x* reaches a bound at ±B, whichever occurs first (Kiani et al., 2008).

When T_s is offered, we assert that the monkey exercises or waives this option based on the probability that a decision based on the sign of $x(t_d)$ will be correct. The model assumes implicit knowledge of the mapping between $\left[x(t_d), t_d\right]$ and this probability. First, we calculate the log likelihood ratio that any $x(t_d)$ will be achieved with motion in the preferred direction (Kiani and Shadlen, 2009):

$$\Lambda_{pref}\left(x(t_d), t_d\right) = \log \frac{\sum_i P\left[x\left(t_d\right) \middle| C_i\right] P\left[C_i\right]}{\sum_i P\left[x\left(t_d\right) \middle| -C_i\right] P\left[-C_i\right]} \qquad \text{(Equation S4)}$$

Preferred choices ($x(t_d) > 0$) are associated with positive $\Lambda_{pref}\left(x(t_d), t_d\right)$. The log likelihood ratio for the null motion direction can be easily calculated by inverting the ratio on the right hand side of Equation S4 (i.e., $\Lambda_{null}\left(x(t_d), t_d\right) = -\Lambda_{pref}\left(x(t_d), t_d\right)$), making Λ_{null} a flipped version of Λ_{pref} around x=0. Because the two directions of motion are equally probable, *a priori*, we can calculate the log posterior odds of a correct choice as $\Lambda_{cor} = \left|\Lambda_{pref}\right|$. Thus, Λ_{cor} is a symmetrical map around zero, and $\Lambda_{cor}\left(0, t_d\right) = 0$. Equation S4 requires estimates of the likelihoods, $P\left(x(t_d), t_d \middle| C_i\right)$. In the present experiment, we do not measure response times, so we lack experimental access to the decision time, t_d, but we can calculate its distribution for each trial type using the model.

A central assumption of the model is that Λ_{cor} is not affected by μS (see discussion of extended model, below). It is a mapping of certainty about evidence, independent of valuation, bias and stopping criterion. To calculate Λ_{cor} we use $R = \alpha C$ and $\sigma^2 = 1$. Therefore, Λ_{cor} depends on only one model parameter, α (Beck et al., 2008; Moreno-Bote, 2010), and the set of motion strengths in the experiment (Kiani and Shadlen, 2009).

As in previous studies, our monkeys exhibited a bias against the preferred direction of the stimulated neurons. This compensatory 'null choice' bias is thought to arise from the subject's attempt to equalize the two choices in the face of an increasing tendency (caused by μS) to choose the preferred direction (Salzman et al., 1992). Although there are several ways to implement such a bias in diffusion models (Ratcliff and McKoon, 2008; van Ravenzwaaij et al., 2012), we find that we can account for it most parsimoniously with an offset (γ) to the drift rate, which approximates a dynamic bias signal (Hanks et al., 2011). Thus for no-μS trials,

$$R_{-\mu S} = \alpha C + \gamma \qquad \text{(Equation S5)}$$

$$\sigma^2_{-\mu S} = 1 \qquad \text{(Equation S6)}$$

This bias is also present on μS trials. In addition, μS can potentially affect both the drift rate and the sensory noise:

$$R_{+\mu S} = \alpha\left(C + \delta_C\right) + \gamma \qquad \text{(Equation S7)}$$

$$\sigma^2_{+\mu S} = 1 + \delta_{\sigma^2} \qquad \text{(Equation S8)}$$

where adjustments to constants are indicated by δ_x using the subscript to indicate the modified term. The term δ_C would cast the effect of μS in units of motion strength.

For each trial, we obtain the probability distribution of the decision variable $P(x,t)$ by numerical solution to the Fokker-Planck equation (Chang and Cooper, 1970) using the definitions of R and σ^2 in Equations S5–S8. The solution establishes the distributions of upper

and lower bound absorption — $P(x=B,t)$ and $P(x=-B,t)$, respectively — and the distribution of x within the bounds (unabsorbed probability, $P(-B<x<B,t)$). These distributions are then used to calculate the expected probabilities of each choice under the model parameters.

On trials when T_s is not shown (T_s^-), the probability of a preferred direction choice is the probability that x has terminated in the upper bound or that it has a final positive value:

$$P_{pref}^{T_s^-}(C,\tau)=\int_0^{\tau}P\left(x=B,t_d\right)dt_d+\int_{x>0}^{x<B}P\left(x,t_d=\tau\right)dx+0.5P\left(x=0,t_d=\tau\right) \qquad \text{(Equation S9)}$$

where τ is the stimulus duration. $P_{null}^{T_s^-}(C,\tau)$ can be simply calculated as $1-P_{pref}^{T_s^-}(C,\tau)$.

On trials when T_s is shown (T_s^+), the probability that the monkey chooses this option depends on a criterion, θ, which is in units of Λ_{cor} (log odds correct). By inverting Λ_{cor} we obtain a criterion in units of the decision variable such that $x_\theta(t)=\left|\Lambda_{cor}^{-1}(\theta,t)\right|$. $x_\theta(t)$ defines the decision variable that makes $\Lambda_{cor}=\theta$ at each moment in time (black curves in Figures 4B–4D). Decision variables in the range $\left[-x_\theta(t),+x_\theta(t)\right]$ lead to sure-bet choices. Decision variables larger than $+x_\theta(t)$ lead to a preferred direction choice and those smaller than $-x_\theta(t)$ lead to a null choice. Thus, the probability of a sure-bet choice is

$$\begin{aligned}P_{sb}^{T_s^+}(C,\tau)=&\int_{-x_\theta(\tau)}^{x_\theta(\tau)}P\left(x,t_d=\tau\right)dx+\\&\int_0^{\tau}P\left(x=B,t_d\right)F\left(\Lambda_{cor}\left(B,t_d\right)\right)dt_d+\\&\int_0^{\tau}P\left(x=-B,t_d\right)F\left(\Lambda_{cor}\left(-B,t_d\right)\right)dt_d\end{aligned} \qquad \text{(Equation S10)}$$

where $F(\Lambda)$ equals 1 if $\Lambda<\theta$ and 0 otherwise. The last two integrals in Equation S10 capture the possibility that a decision that terminates at a bound would lead to a T_s choice — that is, when $\Lambda_{cor}\left(x=\left|B\right|,t_d\right)<\theta$ (e.g., for small B; not the case in Figures 4B–4D). The probability of a

preferred direction choice is given by

$$P_{pref}^{T_s^+}(C,\tau) = \int_{x>x_\theta(\tau)}^{B} P\left(x, t_d = \tau\right) dx + \int_0^\tau P\left(x = B, t_d\right)\left(1 - F\left(\Lambda_{cor}\left(B, t_d\right)\right)\right) dt_d \quad \text{(Equation S11)}$$

and

$$P_{null}^{T_s^+}(C,\tau) = 1 - P_{pref}^{T_s^+}(C,\tau) - P_{sb}^{T_s^+}(C,\tau). \quad \text{(Equation S12)}$$

We used a tiered maximum likelihood fitting procedure to obtain the model parameters. First, we fit the three basic model parameters ($\alpha, B,$ and θ) (Kiani and Shadlen, 2009) and the compensatory bias (γ) by maximizing the joint probability of the observed outcomes (i.e., preferred choice, null choice, or T_s choice when available) for the no-μS trials only (blue filled and open data points in Figure 2B). With these four parameters held fixed, we then estimated δ_C and δ_{σ^2} using only the probability of T_s choices on μS trials (Figure 2B, top, red data points).

Finally, having estimated all the model parameters, we generated *predictions* for the stimulation effect on direction choices (Figure 2B, bottom, red curves) and the improvement in sensitivity on μS trials when T_s was waived (Figure 2B, bottom, solid vs. dashed red curve).

In total, the model has six free parameters, but note that only two of them (δ_C and δ_{σ^2}) are used to explain the effects of μS on T_s choices and no additional parameters are used for the effects of μS on the choice functions. Moreover, our tiered fitting procedure prevents fine tuning of the basic model parameters by the stimulated trials which might have artificially boosted the explanatory power of the model. We repeated the fitting procedure multiple times from different starting points to ensure our results were not driven by local minima. Standard errors (s.e.) for the parameters were calculated using a bootstrap procedure (see Table 1).

In addition to the 'extended' model described below, we considered two alternative nested models to test whether μS affects the sure-bet criterion (θ). In the first, the parameter δ_{σ^2} was set to zero and a new free parameter was added to implement an offset to θ on μS trials (6

free parameters); that is, $\theta_{-\mu S} = \theta$ and $\theta_{+\mu S} = \theta + \delta_\theta$. In the second nested model, δ_{σ^2} was allowed to vary in addition to δ_θ (7 free parameters). These two alternative models were compared with the basic model described above using the Bayesian Information Criterion (BIC), where $\Delta BIC = BIC_{alt} - BIC_{basic}$. The value of ΔBIC was positive for both comparisons (basic vs. 6-parameter alternative, ΔBIC = 130.5; basic vs. 7-parameter alternative, ΔBIC = 10.9), indicating a better fit of the basic model while taking into account the number of free parameters.

Extended model for fitting high-current stimulation data

In the main text, we reported that stimulation of MT/MST with 75 μA current increased the proportion of T_s choices, consistent with lower confidence on these trials. Here we describe a model that explains this and other features of the high-current data, and we compare this model to the simpler one used for the main experiments (see above and Experimental Procedures). This modeling exercise is instructive, and while the key findings are statistically reliable, they should be regarded as tentative because they are based on only eight stimulation sites and relatively few trials (4,483; compare to 53,134 for low-current μS). To obtain a larger data set on this control would have necessitated many fewer trials with low current.

The extended model contains 9 degrees of freedom. To avoid local minima, we fit the model from several random starting points. We adopted the best of these to estimate the parameters and repeated this process in a bootstrap analysis to estimate standard errors for the parameters. That is, the s.e. are the standard deviations of the best fit to a resampled (with replacement) data set, applying the same procedure of multiple random starting points for each iteration. The parameter values and their s.e. are displayed in Table S1.

A summary of the model is as follows. The deterministic drift without μS is identical to the simpler model:

$$R_{-\mu S} = \alpha C + \gamma \qquad \text{(Equation S13)}$$

where α converts motion coherence to the signal component of signal-to-noise and γ is the bias term. As in the simpler model, microstimulation can affect the drift rate as an offset to the motion strength, replacing αC with $\alpha(C+\delta_C)$. However, we also consider the possibility that stimulation spreads to direction columns that supply negative evidence, thereby changing the relationship between C and drift rate. Thus we allow for a change to α:

$$R_{+\mu S} = (\alpha + \delta_\alpha)(C + \delta_C) + \gamma \qquad \text{(Equation S14)}$$

In other words, the model attributes the μS-induced change in sensitivity to a reduction in signal-to-noise as if the population response were less direction selective. Thus, the slope of the choice function is flatter and T_s choice frequency does not fully attenuate with high motion strength (Figures 3, S2A). This is explained by a degradation of signal (δ_α) rather than an increase in noise (δ_{σ^2}), as if high current μS reduces the difference between firing rates of neurons with different direction preferences, as previously suggested (Murasugi et al., 1993). This effect on signal could be achieved by changes to the visual stimulus, albeit using a different type of random dot display than the one we used. Importantly, the offset to α affects only the distribution of the decision variable during μS trials (Equations S9–S11), not the mapping of DV to confidence (Equation S4). The value of δ_α was small but still significantly different from zero for the low-current dataset, suggesting that the decrease in sensitivity (Figure 2B, bottom) was not fully explained by increased noise (δ_{σ^2}). This is expected, however, since even low-current stimulation would occasionally recruit a broader distribution of direction signals than intended.

The extended model also includes a more elaborate description of the noise itself, irrespective of microstimulation. For all trials, we relaxed the constant variance assumption and included a coherence-dependent term:

$$\sigma^2_{-\mu S} = 1 + \beta|C|$$

$$\sigma^2_{+\mu S} = 1 + \delta_{\sigma^2} + \beta|C| \qquad \text{(Equation S15)}$$

The idea that noise scales with motion strength is consistent with the observation that the MT response to preferred and null direction motion is asymmetric. Specifically, an increase in rightward motion coherence causes an increase in the firing rate of cells that prefer rightward which is about 3-fold larger than the decrease in firing rate of cells that prefer leftward (Britten et al., 1993). This implies that the difference in firing rates of pools of MT neurons with opposing direction preferences should should have a variance that is proportional to $|C|$ via multiplication with β. We typically ignore this signal-dependent noise for the sake of simplicity and because the impact happens to be less pronounced for the low-current dataset (Table S1), but including it for the smaller high-current dataset qualitatively improved the fit.

Lastly, we found that the high-current fits could be further improved by defining the boundaries of the choose/waive T_s option with greater flexibility. Recall that this is captured by a single scalar value, θ, in the simpler model used in the main text. Here we allow an asymmetry in this value for preferred and null direction choices (Figure S2B) — i.e., θ_1 and θ_2 — allowing the model to capture a tendency to opt out more often for preferred or null direction decisions. The resulting asymmetry decouples the choice functions on T_s-waived vs. T_s-absent trials (Figure S2A, bottom, solid vs. dashed curves), suggesting an alteration of the mapping between DV and certainty in those sessions in which high current stimulation was present. Note that the values of θ_1 and θ_2 are not nearly as discrepant for low current μS (Table S1), which maintains the approximate symmetry of the mapping as assumed in the simple model in the main text. We also tested a more elaborate model which allowed an offset to θ_1 and θ_2 on μS trials only (not shown), but these offset parameters were not significantly different from zero (high-current data, $P > 0.81$; low-current data, $P > 0.30$, consistent with model comparisons using BIC, mentioned above). Again, however, we are not prepared to draw strong conclusions about this feature of the

data from only a few stimulation sites in one monkey. The primary role of the high-current experiment is to show that it is possible in our paradigm to induce a decrease in confidence (i.e., an increase in T_s choices).

SUPPLEMENTAL REFERENCES

Beck, J.M., Ma, W.J., Kiani, R., Hanks, T., Churchland, A.K., Roitman, J., Shadlen, M.N., Latham, P.E., and Pouget, A. (2008). Probabilistic population codes for Bayesian decision making. Neuron *60*, 1142-1152.

Britten, K.H., Shadlen, M.N., Newsome, W.T., and Movshon, J.A. (1993). Responses of neurons in macaque MT to stochastic motion signals. Vis Neurosci *10*, 1157-1169.

Chang, J.S., and Cooper, G. (1970). A practical difference scheme for Fokker-Planck equations. J Comput Phys *6*, 1-16.

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.

Hof, P.R., and Morrison, J.H. (1995). Neurofilament protein defines regional patterns of cortical organization in the macaque monkey visual system: a quantitative immunohistochemical analysis. J Comp Neurol *352*, 161-186.

Moreno-Bote, R. (2010). Decision confidence and uncertainty in diffusion models with partially correlated neuronal integrators. Neural Computation *22*, 1786-1811.

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

van Ravenzwaaij, D., Mulder, M.J., Tuerlinckx, F., and Wagenmakers, E.J. (2012). Do the dynamics of prior information depend on task context? An analysis of optimal performance and an empirical test. Front Psychol *3*, 132.