Learned Predictions of Error Likelihood in the Anterior Cingulate Cortex

Joshua W. Brown and Todd S. Braver

Science **307**, 1118 (2005);
DOI: 10.1126/science.1105783

6. P. Chen *et al.*, *Nature Cell Biol.* **5**, 422 (2003).
7. M. Classon, E. Harlow, *Nature Rev. Cancer* **2**, 910 (2002).
8. M. M. Lipinski, T. Jacks, *Oncogene* **18**, 7873 (1999).
9. M. Classon, N. Dyson, *Exp. Cell Res.* **264**, 135 (2001).
10. T. Jacks *et al.*, *Nature* **359**, 295 (1992).
11. K. L. Ferguson, R. S. Slack, *Neuroreport* **12**, A55 (2001).
12. D. M. Thomas *et al.*, *Mol. Cell* **8**, 303 (2001).
13. R. S. Slack, H. El-Bizri, J. Wong, D. J. Belliveau, F. D. Miller, *J. Cell Biol.* **140**, 1497 (1998).
14. K. L. Ferguson *et al.*, *EMBO J.* **21**, 3337 (2002).
15. D. MacPherson *et al.*, *Mol. Cell. Biol.* **23**, 1044 (2003).
16. S. Marino, D. Hoogervoorst, S. Brandner, A. Berns, *Development* **130**, 3359 (2003).
17. Z.-Y. Chen, data not shown.
18. M. Vooijs, H. te Riele, M. van der Valk, A. Berns, *Oncogene* **21**, 4635 (2002).
19. F. Liu *et al.*, *Int. J. Dev. Biol.* **48**, 645 (2004).
20. P. Chen, J. E. Johnson, H. Y. Zoghbi, N. Segil, *Development* **129**, 2495 (2002).
21. M. Xiang, W. Q. Gao, T. Hasson, J. J. Shin, *Development* **125**, 3935 (1998).
22. A. L. Woods *et al.*, *Histopathology* **19**, 21 (1991).
23. L. Zheng *et al.*, *Cell* **102**, 377 (2000).
24. R. J. Goodyear *et al.*, *J. Neurosci.* **23**, 9208 (2003).
25. J. Zhang *et al.*, *Nature Genet.* **36**, 351 (2004).
26. J. E. Gale, W. Marcotti, H. J. Kennedy, C. J. Kros, G. P. Richardson, *J. Neurosci.* **21**, 7013 (2001).
27. J. R. Meyers *et al.*, *J. Neurosci.* **23**, 4054 (2003).
28. G. S. Geleoc, J. R. Holt, *Nature Neurosci.* **6**, 1019 (2003).
29. J. Sage *et al.*, *Nature* **424**, 223 (2003).
30. J. R. Holt *et al.*, *J. Neurophysiol.* **81**, 1881 (1999).
31. M. A. Vollrath, R. A. Eatock, *J. Neurophysiol.* **90**, 2676 (2003).

32. We thank L. Goodrich, D. Fekete, M. Holley, and J. Macklis for critical comments on the manuscript. D.P.C. is an Investigator of the Howard Hughes Medical Institute. Supported by NIH grants DC-04546 (Z.-Y.C.), DC-00200 (J.T.C.), and DC-AG20208 (P.W.H.) and by a Pfizer/AFAR Innovations in Aging Research grant (Z.Y.C.).

Supporting Online Material
www.sciencemag.org/cgi/content/full/1106642/DC1
Materials and Methods
Figs. S1 to S6
References

22 October 2004; accepted 3 January 2005
Published online 13 January 2005;
10.1126/science.1106642
Include this information when citing this paper.

Learned Predictions of Error Likelihood in the Anterior Cingulate Cortex

Joshua W. Brown* and Todd S. Braver

The anterior cingulate cortex (ACC) and the related medial wall play a critical role in recruiting cognitive control. Although ACC exhibits selective error and conflict responses, it has been unclear how these develop and become context-specific. With use of a modified stop-signal task, we show from integrated computational neural modeling and neuroimaging studies that ACC learns to predict error likelihood in a given context, even for trials in which there is no error or response conflict. These results support a more general error-likelihood theory of ACC function based on reinforcement learning, of which conflict and error detection are special cases.

Despite remarkable recent advances in cognitive neuroscience, it remains unclear how the brain learns to exert cognitive control over behavior. ACC and neighboring areas in the frontal medial wall play a role in monitoring and controlling goal-directed behavior (*1*–*3*). Error-related negativity (ERN/Ne) (*4*–*6*) and single-unit studies propose that ACC detects errors as discrepancies between actual and intended events (*7*). Alternatively, ACC may detect conflict between mutually incompatible response processes (*8*–*10*) such as incorrect versus correct responses. However, it has not been clear how the ACC learns what constitutes an error or that a given set of responses conflict. We develop here a computational model that demonstrates how ACC might not detect conflict or errors per se but rather more generally represent a prediction of error likelihood (*11*). In particular, the model makes a very specific prediction regarding ACC activity dynamics: The ACC response to a given task condition will be proportional to the perceived likelihood of an error in that condition.

The error-likelihood hypothesis also posits a training signal by which stimulus-specific ACC effects are acquired. This training signal may be dopaminergic. Phasic midbrain dopamine neuron activity is critically involved in reinforcement learning (*12*, *13*), and phasic suppression of dopamine apparently drives the ERN/Ne (*11*). Thus, phasic dopamine suppression occurring in response to errors might serve as a training signal (*14*, *15*) that causes ACC to respond more strongly to contexts in which errors are more frequent.

We formalized the error-likelihood hypothesis as a computational neural model to (i) examine how ACC representations might develop through experience and (ii) explicitly investigate the implications of the hypothesis (*16*). To test the model, we used a change variant (Fig. 1) of the well-known stop-signal paradigm used to examine inhibitory control (*17*). Previous work has demonstrated that stop-signal trials are associated with increased ACC activity in humans (*18*) and related medial frontal areas in other primates (*7*, *19*). The task consisted of conditions associated with high and low error rates crossed with the presence or absence of response conflict. Comparison of correct trials without response conflict across high and low error-rate conditions afforded assessment of error-likelihood effects while controlling for response conflict and errors.

In the computational model (Fig. 2A), simulated ACC neuron units received inputs repre-

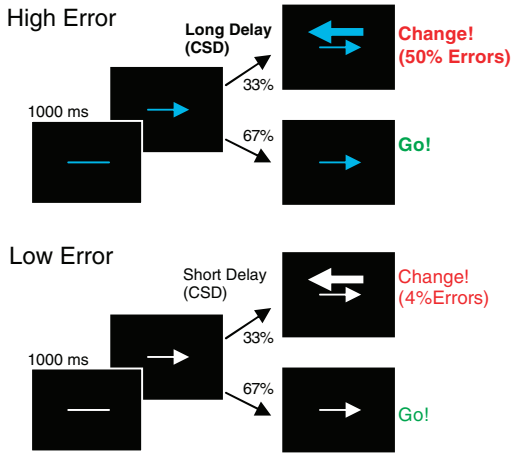

Fig. 1. Change signal task. An initial difficulty cue (plain line) appeared for 1000 ms. For a given trial, cue color (white or blue, equiprobable and randomly intermixed) predicts low versus high error likelihood, respectively. Then a go signal was presented (left- or right-pointing arrow made by extending the plain line to → or ←) that indicated the required button-press response (left index finger for left arrow and right index finger for right arrow). On 33% of the trials, a change (conflict) signal was added to the go signal after a variable CSD relative to go signal onset. This change signal (a second, larger arrow appearing above the first and pointing in the opposite direction) indicated that the response had to be left-right reversed from that indicated by the go signal. Both change and go signals remained visible until a response deadline of 1000 ms after go signal onset. A 500-ms blank interval occurred before the onset of the next trial. All stimuli for a given trial were the same color. Error rates were explicitly set and controlled by dynamically adjusting the CSDs for each error-likelihood condition independently with the use of a staircase algorithm. The low-error condition had shorter CSDs, whereas the high-error condition had longer CSDs, reflecting the well-established positive monotonic relationship between CSD and error rate in stop-signal tasks (*17*).

Department of Psychology, CB 1125, Washington University, St. Louis, MO 63130, USA.

*To whom correspondence should be addressed. E-mail: jwbrown@artsci.wustl.edu

senting stimulus features. Neuroanatomically, there is a lack of evidence for direct sensory inputs to ACC; yet such inputs may arrive either via high-level perceptual pathways or frontal areas that show stimulus specificity (*20*). When errors were committed, ACC units also received an error signal representing phasic suppression of afferent midbrain dopaminergic signals (*12*, *13*). ACC units were adaptive such that the synaptic weight from a given stimulus-specific representation to a given ACC unit increased when both were active and the error signal was also present. Learning within the ACC was competitive. As weights increased from active stimulus-specific inputs to ACC units that were active during an error trial, weights decreased from inactive stimulus units. This constraint, along with lateral excitation between spatially contiguous ACC units, global lateral inhibition among ACC units, and random activity noise, led to a greater number of spatially contiguous ACC units representing conditions reliably associated with errors.

The adaptive error-likelihood model was directly compared against an alternative competing computational model of ACC function (Fig. 2B) in which ACC measures response conflict in terms of the coactivation of incompatible response representations (*21*). In the response conflict model, input synaptic weights to the ACC were fixed rather than adaptive and originated from response rather than stimulus-specific representations.

In both models, activity from ACC excited a control signal that sent persistent, nonspecific inhibition to the response layer. This in turn slowed responses on trials after high ACC activity. Although other kinds of control signals may also be generated by ACC, this control mechanism has been examined in previous modeling work and found sufficient to account for trial-to-trial adjustments in behavioral performance driven by ACC activity fluctuations (*22*).

Simulations conducted with the change-signal task revealed that both the error-likelihood and conflict ACC models showed a similar pattern of behavioral performance, and subsequent experimental testing revealed that both provided very good fits to human behavioral data (*16*). However, despite being fit only to behavioral and not neuroimaging data, the two models showed important differences in the pattern of ACC activity exhibited across task conditions. In the error-likelihood model, ACC activity was greater for change than for go trials [change activity was 0.07 and go activity was 0.05, $F(1,17) = 757$, $P < 0.0001$], as expected (Fig. 2A), but was also greater on high-error change trials than on low-error change trials [high activity was 0.11 and low activity was 0.06, $F(1,17) = 263$, $P < 0.0001$], even when excluding all trials where errors were committed. Even more surprisingly, this error-likelihood effect was also present on correct go trials, which should have no associated conflict [high activity was 0.07 and low activity was 0.03, $F(1,17) = 519$, $P < 0.0001$]. Nevertheless, the error-likelihood model exhibited this effect, because high-error go trials contained a stimulus feature (color) that was associated with an increased likelihood of error commission. Because of the learned associations (embodied as strong synaptic connectivity) between this stimulus feature and ACC units, the presentation of this feature, even in the absence of other features that might cause conflict or errors (i.e., the change signal), was sufficient to increase activity in the ACC. In contrast, the conflict ACC model did not show this pattern of activation (Fig. 2B). Specifically, the conflict model did not exhibit a correct-trial high-error go versus low-error go effect [high activity was 0.13 and low activity was 0.13, $F(1,17) = 0.47$, $P = 0.50$].

Thus, the two models, which instantiate alternative accounts of ACC function, make different predictions regarding the pattern of ACC activity during the change-signal task. In particular, a critical test of the error-likelihood versus conflict-monitoring hypotheses is whether or not the correct, high-error condition go trials yield greater ACC activity than correct, low-error condition go trials. Both of these conditions exclude conflict, change signal delay (CSD), and error effects. Random equiprobable interleaving of high- and low-condition trials further ensures that effects of this comparison could only be driven by current cue color, which predicts error likelihood. We investigated these predictions experimentally in an event-related functional magnetic resonance imaging (fMRI) study of 16 human participants performing the same change-signal task simulated with the models (*16*).

Two regions of interest (ROIs) within the ACC and the neighboring functionally related

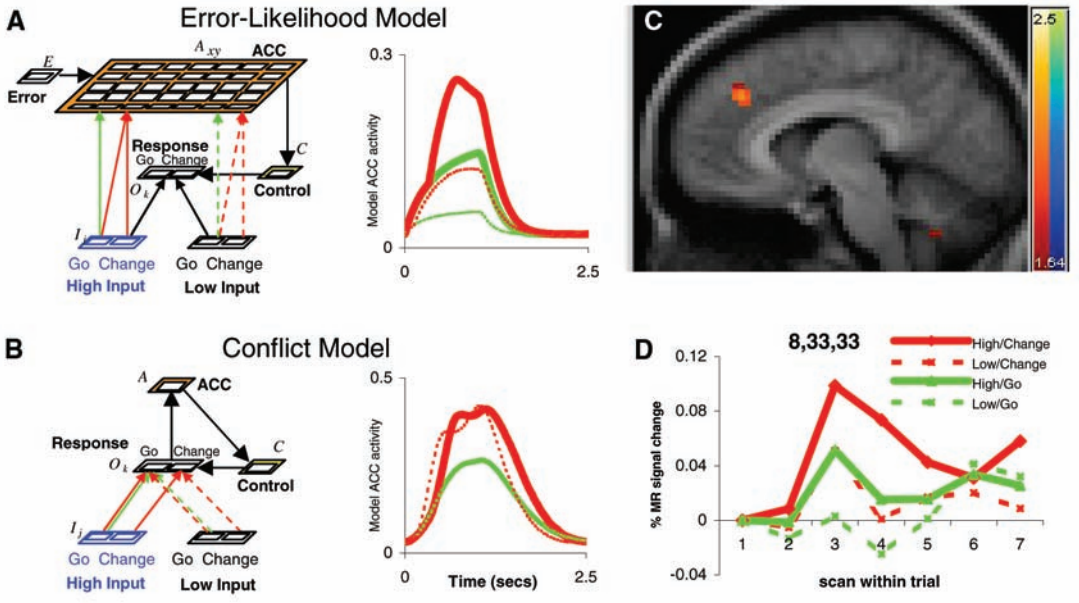

Fig. 2. Error-likelihood effects. Computational models of (**A**) error likelihood and (**B**) conflict detection generate competing predictions regarding effects of high versus low error rate conditions on ACC activity (*16*). Only the error-likelihood model predicts effects of high-error–versus low-error–likelihood conditions in the low-conflict go (thick green versus thin green lines) as well as high-conflict change conditions (thick red versus thin red lines). Model activity is proportional to neural firing rate. (**C**) Region 1 at (8,33,33) in right ACC was identified by a conjunction of several independent statistical tests for significant effects during correct trials. Shown is the z score of high/go minus low/go (restricted to voxels for which $P < 0.05$, one-tailed, uncorrected t test), masked by voxels that also showed a significant effect of both change > go and high/change > low/change (each $P < 0.05$, one-tailed, uncorrected t test). The conjunction analysis mask reduces the sample size considerably to only a few regions, effectively correcting for multiple comparisons while maintaining sensitivity to complex effects. (**D**) Time course of event-related fMRI activation in region 1 (averaged across all voxels) in each task condition. Analysis of the activation-magnitudes (again, averaged across whole region) confirms that all predicted effects are statistically significant [for change > go, $t(15) = 2.52$ and $P < 0.03$; for high/change > low/change, $t(14) = 2.25$ and $P < 0.05$; and for high/go > low/go, $t(14) = 2.17$ and $P < 0.05$]. Time courses shown are those without reaction time effects partially correlated out.

medial wall performance-monitoring areas (*23*) were identified that showed significant effects ($P < 0.05$) for all three of the following contrasts (analyzed by using only correct trials): change greater than go (change > go), high error-likelihood/change greater than low error-likelihood/change (high/change > low/change), and high error-likelihood/go greater than low error-likelihood/go (high/go > low/go). Region 1 (22 voxels; right dorsal ACC; Talairach 8,33,33) (Fig. 2C) overlapped anatomically with an ACC region previously identified to show conflict effects in several go/no go and stop tasks (*18*). Region 2 was found more superiorly (18 voxels; left anterior pre-supplementary motor area (pre-SMA); –2,27,54) with change > go [$t(15) = 2.87$, $P < 0.02$] and high > low effects in both change [$t(14) = 3.52$, $P < 0.004$] and go [$t(14) = 2.56$, $P < 0.03$] conditions (*24*). Critically, ROI analyses (Fig. 2D) confirmed that the effects observed within these regions were highly similar to those observed in the error-likelihood model simulation (Fig. 2A) (*25*). Other ROIs in the right dorsolateral prefrontal cortex and bilateral inferior parietal lobule and cerebellum also showed these same effects (table S1) (*16*) but were not analyzed in detail. Nevertheless, it is possible that these regions implement ACC-driven control signals (*1*), consistent with our computational model.

The error-likelihood hypothesis further postulates that error predictions in the ACC are acquired through task experience via a dopaminergic-training signal that is elicited after error commission. Simulations of the error-likelihood and conflict models again made competing predictions in this regard (Fig. 3, A and B). Specifically, the error-likelihood but not the conflict-detection simulation showed gradually emerging effects of high versus low conditions over the course of a session. fMRI results (Fig. 3C) are consistent with the error-likelihood model (Fig. 3A) but not the conflict model (Fig. 3B). With regard to the origin of training signals driving these learning effects, analysis of fMRI blood oxygen level–dependent (BOLD) responses also revealed increased ACC activity on error trials that differentiated between errors committed in the high and low conditions in a manner consistent with a dopaminergically based training signal (Fig. 4).

The current results support an error-likelihood account of ACC function in which ACC learns to signal, via the magnitude of its activity, the predicted likelihood of an error occurring in response to a given task condition. This proposed function may be highly adaptive, serving as an early-warning system that recruits cognitive control to match its predicted demand. This matching behavior of ACC responses (and subsequent cognitive control) may be seen as a complement to Herrnstein's matching law (*26*). Whereas the matching law predicts increases in neural activity associated with motor responses in proportion to the likelihood of positive reinforcement (*27*), the error-likelihood model predicts that ACC activity will occur in proportion to the likelihood of negative reinforcement. Furthermore, our results are consistent with the idea that a dopaminergic training signal in ACC plays a common role in reinforcement learning and recruitment of cognitive control (*11*).

Conflict and error detection ACC effects are accounted for as special cases of the error-likelihood model. Errors generally occur more frequently with cues that map to conflicting responses compared with cues that map to the same response. Thus, response conflict effects in ACC may reflect a higher error likelihood rather than an explicit computation of response conflict per se. Similarly, previous studies show frontal medial activity in response to errors, even before external feedback (*4*, *28*). Adding a connection in the error-likelihood model from the response to ACC layers might allow ACC to respond selectively to particular combinations of stimuli and internal representations of incorrect response execution, which are highly predictive of undesired consequences. Under these conditions, the model would be expected to show error effects in ACC, even before external feedback (*11*, *28*). Thus, the model (and the associated theoretical framework) provides a tool for understanding how error, conflict, and error-likelihood prediction effects may all be acquired by ACC after continued exposure to task environments. More generally, the results presented here illustrate the benefits of tightly integrating neuroimaging studies with computational modeling, because the two methods together provide a strong basis for hypothesis generation and theory testing regarding the neural mechanisms of cognition.

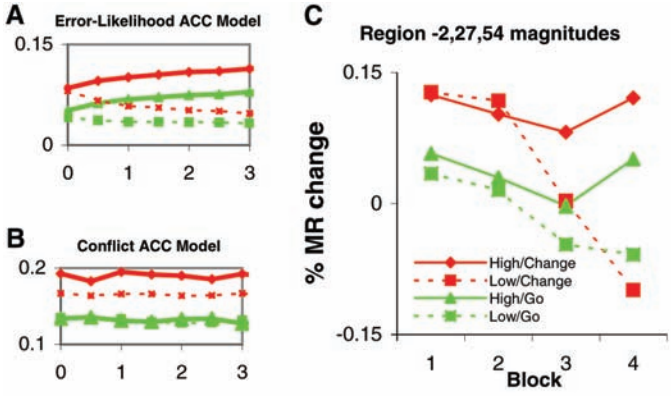

Fig. 3. Learning effects across time. (**A**) The error-likelihood computational model with adaptive ACC inputs predicts an initial effect of change > go and a subsequently growing effect of high > low in each of the change and go conditions. (**B**) The conflict computational model with fixed ACC inputs predicts an effect of high/change > low/change and change > go but no effect of high/go > low/go. None of these effects varied with time. MR, magnetic resonance. (**C**) fMRI results are consistent with the error-likelihood but not the conflict ACC model. During the first two blocks, activity in the anterior pre-SMA region shows effects of change > go but not of high > low, despite the fact that CSDs have equilibrated by the beginning of the second block (fig. S1). By the third and fourth blocks, both the change and go conditions show evidence of discrimination between high- and low-error conditions, and this discrimination eventually dominates the discrimination between change and go trials by the end of the session.

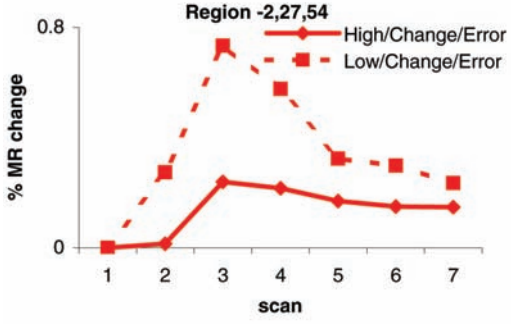

Fig. 4. Error effects consistent with dopaminergic training signals. Error-related activity on change trials is greater for the low-error condition than for the high-error condition [region 2 shown: $t(14) = 3.13$, $P < 0.01$]. When error likelihood is lower, actual errors more highly suppress phasic dopamine cell activity (*29*). According to the error-likelihood hypothesis, dopamine suppression serves as a training signal and increases ACC activity (*11*). Therefore, errors of commission in the low-error condition should be associated with a stronger transient depression of dopamine activity and consequently greater ACC activation than in the high condition. Notably, this pattern is the opposite of that predicted and found for correct trials, where the high condition leads to greater activity than the low condition. We did not model this effect of dopamine signaling in the simulation error signal inputs, but we tested this prediction for the entire ACC region in which voxels showed an effect of change > go ($P < 0.05$). The predicted effect was found in this whole-ACC region [$t(14) = 2.29$, $P < 0.04$], although one subject was excluded for lack of errors in the low/change condition. In general, dopaminergic error signals might also drive plasticity without necessarily exciting target cells, because conflict effects have been found in cells without error responses (*19*).

References and Notes

1. J. G. Kerns *et al.*, *Science* **303**, 1023 (2004).
2. J. D. Schall, V. Stuphorn, J. W. Brown, *Neuron* **36**, 309 (2002).
3. G. Bush, P. Luu, M. I. Posner, *Trends Cogn. Sci.* **4**, 215 (2000).

4. W. J. Gehring, M. G. H. Coles, D. E. Meyer, E. Donchin, *Psychophysiology* **27**, S34 (1990).
5. J. Hohnsbein, M. Falkenstein, J. Hoorman, *J. Psychophysiol.* **3**, 32 (1989).
6. P. S. Bernstein, M. K. Scheffers, M. G. H. Coles, *J. Exp. Psychol. Hum. Percept. Perform.* **21**, 1312 (1995).
7. S. Ito, V. Stuphorn, J. W. Brown, J. D. Schall, *Science* **302**, 120 (2003).
8. C. S. Carter *et al.*, *Science* **280**, 747 (1998).
9. M. M. Botvinick, L. Nystrom, K. Fissel, C. S. Carter, J. D. Cohen, *Nature* **402**, 179 (1999).
10. A. W. MacDonald, J. D. Cohen, V. A. Stenger, C. S. Carter, *Science* **288**, 1835 (2000).
11. C. B. Holroyd, M. G. Coles, *Psychol. Rev.* **109**, 679 (2002).
12. W. Schultz, P. Dayan, P. R. Montague, *Science* **275**, 1593 (1997).
13. J. O'Doherty, P. Dayan, K. Friston, H. Critchley, R. Dolan, *Neuron* **38**, 329 (2003).
14. J. W. Brown, D. Bullock, S. Grossberg, *Neural Networks* **17**, 471 (2004).
15. S. Bao, V. T. Chan, M. M. Merzenich, *Nature* **412**, 79 (2001).
16. Materials and methods are available as supporting material on *Science* Online.
17. G. D. Logan, W. B. Cowan, *Psychol. Rev.* **91**, 295 (1984).
18. K. Rubia *et al.*, *Neuroimage* **13**, 250 (2001).
19. V. Stuphorn, T. L. Taylor, J. D. Schall, *Nature* **408**, 857 (2000).
20. D. Boussaoud, S. P. Wise, *Exp. Brain Res.* **95**, 28 (1993).
21. M. M. Botvinick, T. S. Braver, D. M. Barch, C. S. Carter, J. C. Cohen, *Psychol. Rev.* **108**, 624 (2001).
22. A. D. Jones, R. Cho, L. E. Nystrom, J. D. Cohen, T. S. Braver, *Cogn. Affect. Behav. Neurosci.* **2**, 300 (2002).
23. K. R. Ridderinkhof, M. Ullsperger, E. A. Crone, S. Nieuwenhuis, *Science* **306**, 443 (2004).
24. In addition to the ACC effects, there was also a small but reliable difference in behavioral performance across the two go conditions, with reaction times slightly slower for high/go versus low/go [736 ms versus 710 ms, $F(1,15) = 5.80$, $P < 0.05$]. Because of this pattern, a potential alternative is that the reaction time differences are the causal mechanism underlying differences in ACC activity (if this ACC region was purely sensitive to response-related activity) rather than error-likelihood effects. To rule out this potential confound, ACC activity was reestimated after partially correlating out reaction time as a nuisance covariate. The effects of high/go and low/go remained significant in regions 1 [$t(14) = 2.67$, $P < 0.02$] and 2 [$t(14) = 2.29$, $P < 0.04$].
25. A whole-ACC ROI consisting of all Brodmann areas 24 and 32 voxels that showed significant change > go effects also showed effects of high/change > low/change ($P < 0.04$). The high/go > low/go effect did not reach significance in this ROI ($P = 0.21$), but high/go activity was numerically greater than low/go in 288 of 291 voxels (binomial test, $P < 0.00001$). High and low activities also apparently diverged with training as expected, but this did not reach significance.
26. R. Herrnstein, *J. Exp. Anal. Behav.* **4**, 267 (1961).
27. L. P. Sugrue, G. S. Corrado, W. T. Newsome, *Science* **304**, 1782 (2004).
28. C. B. Holroyd *et al.*, *Nat. Neurosci.* **7**, 497 (2004).
29. C. D. Fiorillo, P. N. Tobler, W. Schultz, *Science* **299**, 1898 (2003).
30. The authors thank C. Hoyer for help with data collection and C. Holroyd, J. Reynolds, A. Schaeffer, J. Schall, and V. Stuphorn for helpful comments. Supported by Office of Naval Research N00014-00-1-0715. Partial funding support was also received by the Director of Central Intelligence's Intelligence Technology Innovation Center. The authors declare that they have no competing financial interests.

Supporting Online Material
www.sciencemag.org/cgi/content/full/307/5712/1118/DC1
Materials and Methods
Fig. S1
Table S1

28 September 2004; accepted 27 December 2004
10.1126/science.1105783

Flexible Control of Mutual Inhibition: A Neural Model of Two-Interval Discrimination

Christian K. Machens,[1] Ranulfo Romo,[2] Carlos D. Brody[1]*

Networks adapt to environmental demands by switching between distinct dynamical behaviors. The activity of frontal-lobe neurons during two-interval discrimination tasks is an example of these adaptable dynamics. Subjects first perceive a stimulus, then hold it in working memory, and finally make a decision by comparing it with a second stimulus. We present a simple mutual-inhibition network model that captures all three task phases within a single framework. The model integrates both working memory and decision making because its dynamical properties are easily controlled without changing its connectivity. Mutual inhibition between nonlinear units is a useful design motif for networks that must display multiple behaviors.

In our daily lives, our minds can flit from thought to thought with remarkable speed and flexibility (*1*). A simplified task that requires rapid shifts between different mental actions is known as two-interval discrimination (two stimuli separated by a time interval; Fig. 1A). Subjects must first perceive a brief stimulus, called f1, maintain it in working memory for several seconds, and then compare it with a brief second stimulus, called f2, to immediately decide which of the two stimuli was larger. The task requires both working memory and decision making, interfacing between the two in a rapid switch from one to the other.

The biophysical mechanisms underlying the performance of this task remain unknown. Spiking neural-network models, built to serve as mechanistic accounts of cognitive neural activity, have focused so far on only single cognitive processes (*2*–*8*). Few models (*9*, *10*), and no spiking network models, have addressed the question of how more than one computation and dynamic can be implemented in a single network. Yet cognitive acts typically require more than one type of computation. Many cognitive psychology models do integrate multiple processes, but do not address biophysical mechanisms (*11*). On the basis of recent neurophysiological data (Fig. 1) (*12*–*17*), we use a nonlinear dynamical systems approach (*18*–*21*) to design a simple and testable spiking-neuron model of two-interval discrimination. The model integrates three key processes into a single framework that proposes mechanistic links between the different processes, as well as between biophysical properties and neural and behavioral phenomena. These processes are fast initial loading of stimulus f1 into working memory, slow maintenance of working memory, and fast decision making.

Figure 1, C and D show the firing rates of two prefrontal cortical (PFC) neurons recorded from previously trained macaque monkeys while they performed a two-interval discrimination task in which f1 and f2 were the frequencies of mechanical vibrations applied to the tip of a finger (*12*, *16*, *22*). The dynamics of the activity of these neurons depends strongly on the phase of the task. During the loading of f1 into working memory, there is a rapid flow to an f1-dependent firing rate. During the maintenance of f1 in working memory, there is a long-lasting persistence of f1-dependent firing rates, despite the absence of the stimulus. During the comparison/decision phase, upon presentation of the second stimulus f2, the firing rates quickly segregate into one of two categories, depending on the monkey's subsequent choice of a "yes" or "no" push-button answer to the question, "Is f1 greater than f2?" Responses similar to these PFC responses are also found in ventral (*17*) and medial (*14*) premotor cortices. For brevity, here we will refer collectively to these three areas as "frontal lobe areas." We highlight two aspects of the frontal lobe data. First, signals are encoded in complementary sets of roughly equal numbers of neurons (*12*, *14*, *17*). One set is composed of "plus" neurons, defined as neurons with a delay-period firing rate that is a monotonically increasing function of f1 (Fig. 1C). Plus neurons typically fire the most for "yes" decisions after presentation of f2. The complementary set are "minus" neurons, defined as those which have delay period firing rates that are monotonically decreasing functions of f1, and fire the most for "no" decisions (Fig. 1D). Because higher f2 stimuli are more likely to lead to "no" decisions; plus neurons are excited by high f1 stimuli but inhibited by high f2 stimuli. The converse occurs for minus neurons

[1]Cold Spring Harbor Laboratory, 1 Bungtown Road, Cold Spring Harbor, NY 11724, USA. [2]Instituto de Fisiología Celular, Universidad Nacional Autónoma de México, 04510 México, D.F., México.

*To whom correspondence should be addressed. E-mail: brody@cshl.edu