A computational model of anterior cingulate function in speeded response tasks: Effects of frequency, sequence, and conflict

ANDREW D. JONES
Washington University, St. Louis, Missouri

RAYMOND Y. CHO
Princeton University, Princeton, New Jersey
and Pittsburgh University, Pittsburgh, Pennsylvania

LEIGH E. NYSTROM
Princeton University, Princeton, New Jersey

JONATHAN D. COHEN
Princeton University, Princeton, New Jersey
and Pittsburgh University, Pittsburgh, Pennsylvania

and

TODD S. BRAVER
Washington University, St. Louis, Missouri

A growing body of evidence from functional neuroimaging and computational modeling studies indicates that the anterior cingulate cortex (ACC) detects the presence of response conflict and conveys this information to other brain regions, enabling subsequent adjustments in cognitive control. The present study examined previous empirical findings of increased ACC for low-frequency stimuli across three distinct speeded response tasks (two-alternative forced choice, go/no-go, and oddball). Simulations conducted in a neural network model incorporating sequential priming mechanisms (developed in Cho et al., 2002) confirmed that a computational measure of response conflict was higher on low-frequency trials across all three tasks. In addition, the model captured detailed aspects of behavioral reaction time and accuracy data, predicted the dynamics of ACC activity related to trial sequence effects, and provided evidence for the functional role of conflict information in performance monitoring and optimization. The results indicate that the conflict-monitoring hypothesis, augmented by mechanisms for encoding stimulus history, can explain key phenomena associated with performance in sequential speeded response tasks.

With the advent of neuroimaging, it has become clear that the anterior cingulate cortex (ACC) is a brain region that is critically involved in cognitive control (Bush, Luu, & Posner, 2000; Carter, Botvinick, & Cohen, 1999; Paus, Koski, Caramanos, & Westbury, 1998; Posner & DiGirolamo, 1998). The exact role of the ACC in cognition is still unknown, but an emerging view is that ACC activation (specifically, within the caudal portion) is directly related to the degree of *response conflict* present in a given cognitive task (Botvinick, Braver, Barch, Carter, & Cohen, 2001; Carter et al., 1998). Response conflict arises because of either physical or task-induced constraints on response generation. It is not possible for certain responses to occur simultaneously. For instance, it is not possible to say the word "green" while at the same time saying the word "red." Conflict can be thought of as the simultaneous coactivation of incompatible responses, such as trying to name the ink color of the word "red" written in green ink in the classic Stroop task. It is difficult to perform such a task because certain conditions generate a tendency to produce an inappropriate response (owing to learned associations or innate biases) in addition to the situationally appropriate one. Cognitive control processes must be used to help mediate performance by overcoming these sources of conflict. A growing neuroimaging literature has demonstrated that the

ACC is reliably activated across a wide range of tasks that involve response conflict, such as the Stroop (Carter et al., 2000), the Eriksen (Botvinick, Nystrom, Fissel, Carter, & Cohen, 1999; Casey et al., 2000), and the go/no-go (Casey et al., 1997) tasks.

In our previous work, we have proposed a specific mechanistic hypothesis as to why the ACC is activated under such response conflict situations (Botvinick et al., 2001; Carter et al., 1998). Specifically, we have hypothesized that the function of the ACC is to provide an on-line index of the degree of response conflict. By doing so, the ACC can convey this information to other systems (such as the prefrontal cortex) that are more directly involved in the implementation of control strategies (i.e., active maintenance of goal-relevant information, attentional selection, response priming). In a series of computer simulation studies, we have demonstrated that this model can account for a variety of empirical phenomena related to ACC activity and behavioral performance in response conflict tasks (Botvinick et al., 2001). Furthermore, a number of recent neuroimaging studies have directly confirmed predictions of the model with regard to activity in the ACC and other regions, such as the dorsolateral PFC (Barch et al., 2001; Barch, Braver, Sabb, & Noll, 2000; Botvinick et al., 1999; Carter et al., 2000; Casey et al., 2000; MacDonald, Cohen, Stenger, & Carter, 2000).

In the present study, we extended the ACC model to account for a more detailed set of behavioral and imaging data. In particular, we examined three different task paradigms: go/no-go, oddball, and two-alternative forced choice (2AFC). On the surface, these three different paradigms appear to be qualitatively distinct in terms of the cognitive demands and processes involved and their relationship to ACC activity. However, we suggest that the tasks tap a common factor, stimulus frequency, that modulates the degree of conflict present on any given trial. Furthermore, this frequency effect on conflict operates in terms of both global stimulus frequency and local sequential history. In order to test this idea, however, existing implementations of the conflict model had to be augmented so as to be responsive to the effects of stimulus frequency on processing. Here, we did so by incorporating the mechanisms for encoding stimulus history developed in Cho et al. (2002). We then tested whether the conflict model, when augmented with mechanisms responsive to stimulus frequency, could capture both the commonalities and the differences between the go/no-go, oddball, and 2AFC tasks in terms of both ACC activity and behavioral performance. Moreover, we investigated whether the model could provide evidence for adjustments in cognitive control strategies that directly result from response conflict. Below, we will discuss the theoretical motivations and empirical data that form the background for the present work.

Conflict and ACC Activity in Speeded Response Tasks

A primary assumption motivating the present study was that conflict arises whenever infrequent responses are re-quired, especially when these occur in the context of making stereotyped or habitual responses. Habitual responses (such as word naming in the Stroop task) have a strong stimulus–response mapping, whereas frequent responses (such as a left buttonpress in a 2AFC with asymmetric probabilities) are primed because of recent frequent engagement. Both situations generate a prepotent response. When the low-frequency response is to be executed, it must compete with and eventually overcome activation of the prepotent response. The competition between the two response pathways is expected to result in processing conflict, which should be reflected in increased ACC activity, even during the performance of otherwise simple cognitive tasks.

In recent neuroimaging work, we directly confirmed this prediction, by demonstrating that manipulations of frequency lead to reliable modulations of ACC activity (Braver, Barch, Gray, Molfese, & Snyder, 2001). Specifically, we observed greater ACC activity for low- versus high-frequency stimuli, but equivalent activity when two different stimuli were of equal frequency. These frequency effects on ACC activity were found in each of three different tasks. The first, 2AFC, required subjects to select one of two responses on the basis of the category of the present stimulus, which varied in frequency across different conditions (low, equal, and high). The other two tasks were the go/no-go and the oddball. In both of these tasks, subjects responded to one stimulus category but withheld responding to the other category. Thus, one distinction between the three tasks was whether two responses (2AFC) or only one response (go/no-go and oddball) was required for performance. The relative frequency of each stimulus category was the task dimension (and the only one in our study) that differentiated the go/no-go and the oddball tasks. In the go/no-go, the *go* stimulus was of high frequency, whereas the *no-go* stimulus was of low frequency. In the oddball, the *go* stimulus (referred to as the *target*) was of high frequency, and the *no-go* stimulus (*nontarget*) was of low frequency. For completeness, we also examined a final go/no-go/oddball hybrid condition, in which the two stimulus conditions were of equal frequency.

A key finding in Braver et al.'s (2001) study was that the frequency effects on ACC activity (greater activity for low- versus high-frequency stimuli) were equivalent across all tasks. This result suggests that the tasks do not differ in terms of the role of frequency in eliciting response conflict (as indexed by ACC activity). However, the fact that the go/no-go and oddball tasks involve the generation or inhibition of a single response, rather than selection between two alternative responses, appears to create problems for the conflict theory of ACC function. Specifically, conflict theory suggests that the ACC detects conflict between *different* response tendencies. It is not clear what the source of conflict could be when there is only a single response option. Moreover, the apparent qualitative distinctions between the tasks seem to be reinforced by their different behavioral performance profiles. Go/no-go tasks are typically associated with high rates of false alarms on no-go trials, whereas targets are only infrequently missed
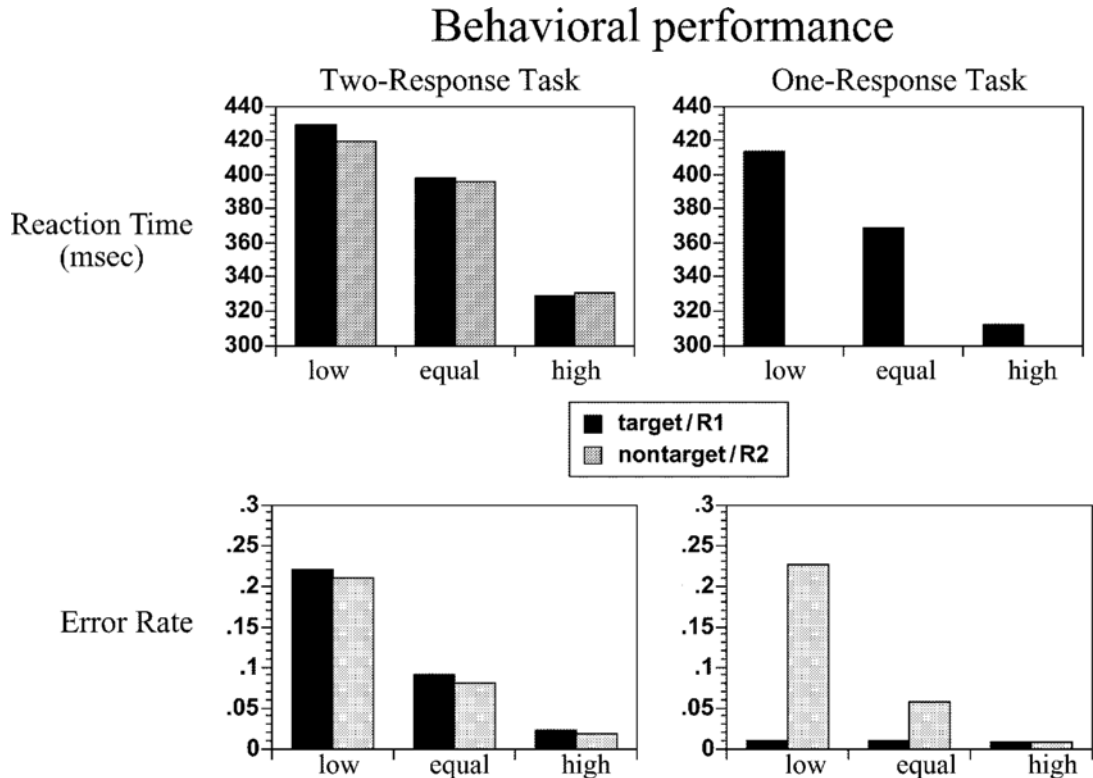
Behavioral performance

Figure 1. Behavioral performance data. Reaction time (upper graphs) and error rates (lower graphs) from both two-response (two-alternative forced choice; left panel) and one-response (go/no-go and oddball; right panel) speeded response tasks. In general, low-frequency stimuli are associated with slower reaction times and more errors. However, there is nearly perfect accuracy for the target stimulus in the one-response conditions, regardless of frequency. Note that the target/R1 stimulus is low frequency when the nontarget/R2 stimulus is high frequency. Thus the dark bar above the label "low" is taken from the same experiment as the light bar above the label "high," and vice versa. Both data points above the label "equal" are taken from the same experimental condition.

in oddball tasks. Figure 1 shows an example of such behavioral performance distinctions across the three tasks (when matched on response frequency). In many cognitive tasks, the degree of response conflict is thought to directly track with behavioral performance (e.g., reaction time in incongruent vs. neutral Stroop trials). It is not clear how to reconcile the obvious behavioral differences between go/no-go and oddball tasks with the notion that they elicit equivalent levels of response conflict under equivalent frequency conditions.

The primary goal of the present paper was to test whether the conflict theory of ACC function can simultaneously capture the equivalence of go/no-go, oddball, and 2AFC tasks in terms of conflict (i.e., ACC activity) while also explaining why the tasks might be dissimilar in terms of behavior. This question is one that can be usefully addressed with a computational modeling approach. Specifically, computational modeling provides a means for explicitly testing whether a hypothesized cognitive or neural mechanism can adequately capture the complexities and constraints of the relevant empirical phenomena. If the model is successful at capturing the desired phenomena, this

adds support for its theoretical sufficiency. In contrast, failures to adequately capture phenomena can be used as evidence for rejection of the theory—at least in its current form. In the present study, we simulated performance of speeded tasks involving both selection between two responses (i.e., 2AFC) and generation versus suppression of a single response (i.e., go/no-go and oddball). We examined whether a single model could account for both the characteristics of ACC activity and behavioral performance in these tasks.

A second goal of the study was to provide an even stronger test of the model by examining whether conflict serves a functionally relevant role in performance. According to the conflict theory, when conflict is high, more control is needed, and when conflict is low, less control is needed. The system is postulated to function as a closed loop, whereby the presence of conflict (as indexed by ACC activity) serves to engage control systems, which then act to reduce the conflict experienced in future situations. In previous computational modeling work, Botvinick et al. (2001) showed that this conflict-control loop, when incorporated into existing computational models of cog-

nitive tasks (e.g., Stroop or Eriksen), enabled these models to successfully capture a number of empirical phenomena associated with cognitive control processes.

Although conflict may serve a functional role in tasks that more clearly rely on the involvement of cognitive control, it is not as clear what functional significance conflict may have in simple speeded response tasks, such as the ones we are concerned with here. One possibility is that subjects might use the ACC-mediated conflict signal as a means for adjusting a speed–accuracy tradeoff function (SATF) so that performance is optimized to the current task conditions. Adjustments to SATFs have been commonly postulated in the literature as a strategic method for optimizing performance in speeded response tasks (Luce, 1986; Rabbitt & Vyas, 1970; Smith & Brewer, 1995) and so might be a critical function of cognitive control. Moreover, there is empirical evidence to support the idea of performance adjustments being made on line during such tasks. Rabbitt (1966) demonstrated that subjects reliably slow down after making an error, whereas Laming (1968) found that subjects become more accurate as well as slower in these same situations. Thus, these findings suggest that subjects opt to decrease speed and increase accuracy after making errors. The conflict theory goes beyond this interpretation to suggest that speed–accuracy adjustments may not be specific to the occurrence of an error per se but, rather, to the experience of a high degree of response conflict. A key principle of the conflict model is that errors are typically associated with high degrees of conflict. In particular, Botvinick et al. (2001) demonstrated that the conflict model could account for error-related performance adjustments solely as a function of the conflict-control loop mechanism, and Yeung, Botvinick, and Cohen (2002) have proposed that conflict monitoring may actually provide a reliable mechanism for error detection.

In the present study, we extended Botvinick et al.'s (2001) simulations by testing whether there are performance adjustments made by subjects in response to high-conflict trials, even when this occurs in the absence of an error. Speeded response paradigms provided an excellent testbed for examining this issue, given that our prior data had indicated that the ACC is activated by low-frequency events even on trials on which no errors are made (Braver et al., 2001). If the ACC activity on these trials is hypothesized to represent the presence of conflict, this conflict should lead to a subsequent adjustment in control (i.e., a shift in the SATF) that would be manifest in behavioral performance. We were specifically interested in determining whether the conflict model could be used to provide evidence for these performance adjustments.

The modeling approach we took was to build on previous computational models developed to examine different aspects of speeded response task performance. Usher and McClelland (2001) implemented a general computational model within the neural network framework to examine the psychological and neural mechanisms underlying the dynamics of choice discrimination. This model was found to be highly successful at capturing a wide range of reaction

time (RT) data associated with speeded response performance. As was mentioned previously, Botvinick et al. (2001) found that this model, when augmented with a conflict-control loop, could capture performance adjustments related to error commission. Most recently, we (Cho et al., 2002) have extended the work of Usher and McClelland to incorporate effects related to sequences of stimuli. In particular, there is a large literature suggesting that a substantial component of variance in two-alternative choice discrimination performance is related to the specific sequential history of trials (Kirby, 1980; Laming, 1968; Remington, 1969; Soetens, Boer, & Hueting, 1985). We were able to account for these data by incorporating priming mechanisms into the basic Usher and McClelland model that are sensitive to both repetitions and alternations in stimulus sequences. The present model extends this previous work by adding the conflict-control loop (Botvinick et al., 2001) into the Cho et al. model. In addition, we modified the model to perform both two-response (i.e., 2AFC) and one-response (i.e., go/no-go and oddball) discrimination tasks. We hypothesized that the priming mechanisms in the model would produce modulations in conflict related to stimulus sequential history and that these modulations would account for the frequency effects on ACC activity observed in the imaging data. Moreover, we tested whether these sequence-related modulations in conflict would lead to adjustments in behavioral performance that could be captured by the presence of the conflict-control loop in the model.

METHOD

The neuroimaging data used for the present study were originally reported in Braver et al. (2001). The behavioral data were collected in the course of extensive pilot testing for that project. We will briefly summarize the experimental design and neuroimaging analysis procedures used to collect this data (see Braver et al., 2001, for further details) and then will focus on the simulations

Experimental Method

Subjects. There were 65 subjects (47 females and 18 males; mean age, 19.3 years; range, 17–25 years) in the behavioral study and 14 subjects (5 males and 9 females; mean age, 22.9 years; range, 18–27 years) in the neuroimaging study. All the subjects in the neuroimaging study were right-handed and neurologically normal. The subjects received either course credit or a monetary reimbursement for their time. Informed consent was acquired according to the guidelines set by the Washington University Human Studies Committee.

Task Design and Procedure. The subjects performed two types of speeded response tasks: one-response (go/no-go/oddball) and two-response (2AFC). In all the tasks, the subjects were presented with a series of uppercase characters in the center of a computer screen (250-msec duration, 1,000-msec interstimulus interval). Furthermore, in all the tasks, the subjects were told that there were two categories of stimuli. In the one-response task conditions, the subjects were asked to generate a buttonpress in response to target stimuli, but not for nontarget stimuli. In the two-response conditions, the subjects were to make an index finger buttonpress (R1) for one category of stimuli and to make a middle finger buttonpress (R2) for the other category of stimuli. Target/R1 stimuli consisted of upper case letters (except for X) and the nontarget/R2 stimulus was the single digit "5" or vice versa. This clearly led to an asymmetry in the

size of target/R1 versus nontarget/R2 categories (25 vs. 1 item). However, the mapping of stimuli to target/R1 and nontarget/R2 categories was counterbalanced across subjects. In the neuroimaging study, the same basic category structure was retained, except that only letters were used as stimuli, with the two stimulus categories being the letter "X" or the remaining letters. The subjects were asked to respond as quickly and accurately as possible.

Each version of the task (one-response, two-response) had three frequency conditions: low-frequency target/R1 (target = 17%), equal-frequency (target = 50%), and low-frequency nontarget/R2 (target = 83%). This made a total of six conditions: (1) one-response low-frequency target (i.e., oddball), (2) one-response low-frequency non-target (i.e., go/no-go), (3) one-response equal frequency, (4) two-response equal frequency, (5) two-response low-frequency R1, and (6) two-response low-frequency R2. Note that Conditions 5 and 6 were essentially identical because of the symmetry of the two-response task and counterbalancing of stimulus categories. Each subject performed two runs of the six conditions, with the order of runs and the specific sequence of stimuli counterbalanced across subjects. Each run consisted of 150 trials. RT and accuracy data were tabulated by computer.

Neuroimaging analysis. A 10-mm spherical region of interest (ROI) within the ACC was generated on the basis of a recent meta-analysis of manual response conflict tasks (Barch et al., 2001). Rapid event-related analysis methods were used to determine ACC activity in response to different stimulus types. ACC activation was expressed in terms of the difference between the low- and the high-frequency responses; for equal-frequency conditions, one stimulus was arbitrarily chosen as the baseline. The peak amplitude of this difference was used as the comparison measure with simulation data.

Simulation Method

Model architecture. As was mentioned above, the basic architecture was developed in Usher and McClelland (2001) and was ex-

tended by Botvinick et al. (2001) and Cho et al. (2002). In the absence of noise, the basic architecture is essentially equivalent to McClelland's cascade model (McClelland, 1979). Our primary modification was to implement an additional response layer to capture the distinction between one- and two-response task conditions, as described below. The network consisted of a stimulus layer, a response decision layer, a response execution layer, and a strategic priming unit (see Figure 2). Conflict was measured at the response decision layer and influenced processing (via a feedback loop through the strategic priming unit) at both the decision and the execution layers of the model. There were two inputs connected in a one-to-one fashion with two decision units in the decision layer. The decision layer units competed with one another via lateral inhibition. Each decision layer unit was, in turn, connected to a response execution unit in a one-to-one fashion. The execution units also competed with each other via lateral inhibition. When the activity level of one of the execution units reached some threshold value, the network was judged to have made an overt response.

The primary difference between the one-response version of the model and the two-response version was that in the one-response model, only one of the decision layer units corresponding to the non-target (no-go) response representation was connected to a response execution unit. The unconnected decision unit could become active and suppress activity in the other decision unit, but activation of the unconnected decision unit did not lead to the execution of an overt response. In this case, the response execution layer still contained two units, but because one of the units did not receive any input from the decision layer, it was given a small baseline activity level. The baseline activity of this unit was high enough to provide some competition with the target execution unit but was low enough that the unit never exceeded its activation threshold. All of the processing units in the decision and execution layers received input from a strategic priming unit. This priming unit was able to influence the global state of the network by altering the baseline activity of the decision and execution units (see the Priming Mechanisms section).
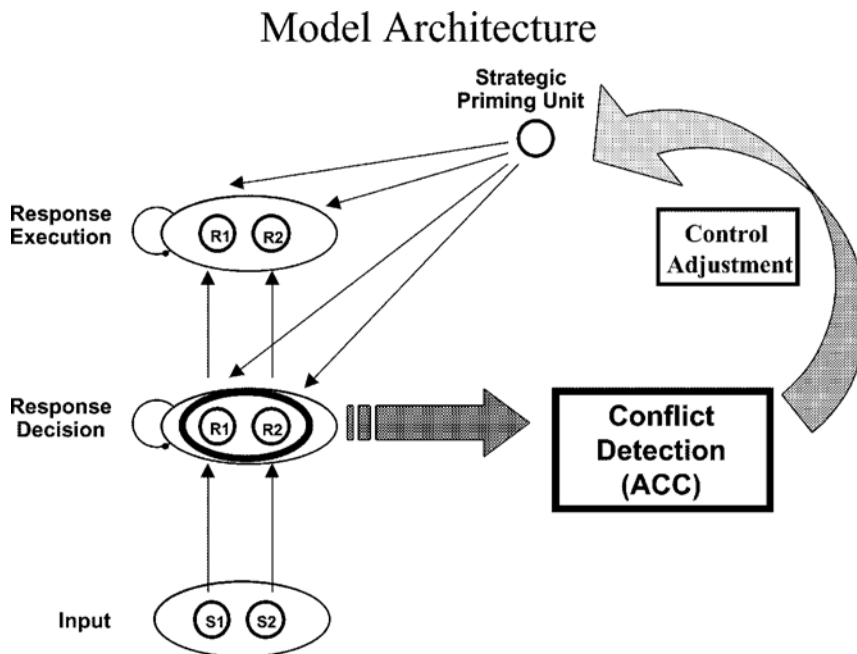

Figure 2. Schematic diagram of model architecture. The network consisted of three feed-forward layers with lateral inhibitory connections at the response decision and response execution layers. Conflict was measured at the decision layer, and modulated activity in the decision and execution layers via a strategic priming unit. Repetition priming and alternation priming (not shown) were also applied to the decision and execution layers.

Processing. Activation dynamics in the simulation were based on those in Usher and McClelland (2001) and Cho et al. (2002). Specific details of the equations governing processing and model parameter selection are described in the Appendix. Units in the model compute their output activation on the basis of integration of information from several sources: excitatory external input (either stimulus input or that from a previous layer), lateral inhibition (from competing units within the same layer), and priming input (both strategic priming and sequential priming).

Each trial consisted of a preparatory phase and a stimulus-processing phase. During the preparatory phase, unit activities were allowed to settle without any external inputs. The preparatory cycle period was intended to establish a baseline level of activity in each processing unit before the presentation of the stimulus and approximated the response–stimulus interval (RSI) typically varied in speeded response experiments (Soetens et al., 1985). Differentially higher baseline activity in one response channel (because of sequential priming or noise) represented a bias to produce that response. Overall, higher or lower baseline activity in each unit (because of strategic priming) represented a global "readiness to respond" in the network. Priming input was present throughout the preparatory-cycle period and during the first portion of the stimulus-processing phase.

During the stimulus-processing phase, the activity of the input units was clamped at a specific pattern corresponding to the present stimulus. Presentation of a particular stimulus always involved some amount of activity in the unit corresponding to the other stimulus (see the Appendix for details). This was intended to encode a degree of stimulus ambiguity and also introduced activity into the incorrect response channel, which provided a means for the network to produce errors when combined with the effects of priming and noise. The input units were transiently activated and then returned to baseline for the duration of the trial, during which activity in the rest of the network was allowed to settle. The model was judged to have made a response when activation of one of the response execution units reached a threshold value. RT was encoded as the number of cycles necessary to make a response after initial presentation of the stimulus. Note that in the one-response version of the model, no overt responses were made for nontarget stimuli, so RT was not a meaningful quantity in this case.

Conflict. Conflict was computed as in previous simulations (Botvinick et al., 2001) and corresponded to a simplified form of the well-known Hopfield energy measure (Hopfield, 1982). In particular, the conflict value for a given trial reflected the simultaneous coactivation of both response decision layer units. Conflict was low when the activity of one decision unit was close to zero, regardless of the activation of the other decision unit (a very small number times a large number is still small). In contrast, conflict was highest when both decision units were very active.

Priming mechanisms. In addition to the standard inputs to each unit, additional sources of input were provided as priming-related activity (see the Appendix for technical details). One type of priming activity was directly modulated by the local sequential history of trials and has been described in Cho et al. (2002). This priming came in two forms: (1) repetition priming (facilitation of a given stimulus after that stimulus had previously been observed) and (2) alternation priming (facilitation of the stimulus opposite to that last observed, following a previous alternation). The particular type of repetition and alternation priming mechanisms used was the combination that provided the best fit to empirical data taken from multiple data sets (Cho et al., 2002).

The second type of priming was due to input from the strategic priming unit. The strategic priming input governed the global preparation of the network for input by changing the baseline activity of units in the manner described by Botvinick et al. (2001). This change was nonspecific (as contrasted to the repetition and alternation priming mechanisms) in that it affected all units equivalently. With lower strategic priming, each unit had a lower level of baseline activity,

which meant that it took a longer period for a response to be made (i.e., slower RT). In addition, a low level of baseline activity also meant that the incorrect response execution unit was less likely to cross its threshold value, decreasing the probability that the network made an error (i.e., higher accuracy). Thus, low strategic priming is associated with more "controlled" behavior. On the other hand, high strategic priming led to more "automatic" behavior (i.e., faster responding and higher error rates, especially for stimuli that were unexpected on the basis of sequential history). The role of the conflict detector was to increase control by decreasing strategic priming in response to conflict. The magnitude of the control adjustment (change in strategic priming) on any given trial was not based simply on conflict from the previous trial. Rather, it was influenced by the conflict from a number of preceding trials, so that each trial backward in the sequence had a successively smaller effect on the present control adjustment.

Certain parameters of the model were adjusted to increase the fit of the model to the behavioral data in preliminary simulations (see the Appendix for details). However, it is important to note that these parameters were fixed prior to the analyses described below and, moreover, were set without regard to effects related to conflict.

Data Analysis

The model was presented with 65 unique sequences of stimuli for each condition, exactly matching the sequences of stimuli presented to each of the 65 subjects in the behavioral study. Each block of 150 trials was treated as a different condition, and each set of six blocks was treated as a different subject for the purpose of statistical comparison against the empirical data. Three sets of analyses were conducted that probed the model at successively finer grains of detail. The first analysis examined how well the model captured behavioral performance and neuroimaging data of ACC activity across tasks as a function of global (i.e., blockwise) stimulus frequency. The second analysis examined the role of local stimulus sequential history in the model on both conflict and behavioral performance, comparing both one-response and two-response tasks. The final analysis examined how well the model captured the trial-by-trial performance of each subject, specifically in terms of the contribution of conflict-driven control adjustments to behavioral performance.

Frequency analysis. To determine the effect of response frequency on performance, mean RT and percentage of error were calculated across subjects (both actual and simulated) for each frequency condition (low-frequency target, equal frequency, and low-frequency nontarget) for each task (one-response and two-response) and for each stimulus (target and nontarget). Only RTs from correct response trials were included in this analysis. Mean simulated conflict for each stimulus was also calculated as a function of frequency and task condition for comparison against imaging data from the ACC (both correct and error trials were used here, in order to match the analysis procedure used in Braver et al., 2001).

Sequential history analysis. Analyses of data were conducted after categorizing each stimulus as a function of the previous five trials, consistent with previous approaches in the empirical literature (Kirby, 1980; Laming, 1968; Remington, 1969; Soetens et al., 1985) and as we have done in our previous simulations (Cho et al., 2002). The categorization scheme coded each stimulus as a repetition (R) or an alternation (A) of the previous stimulus, producing 16 categories. Thus, the sequence of stimuli XXYXX would be characterized as RAAR (repetition, alternation, alternation, repetition). We previously demonstrated that the repetition and alternation priming mechanisms in the model provide a successful fit to the behavioral data in two-response tasks (Cho et al., 2002). In the present study, we performed an additional analysis to compare the fit of the new model (which was slightly modified from the Cho et al. model as a result of an additional layer) across the two-response and the one-response tasks.

We conducted the sequential history analysis only on data from the equal response frequency conditions. This ensured that there were

equal numbers of trials, on average, contributing to each sequential category. Because model parameters were not modified on an individual subject basis, we eliminated individual-difference effects in baseline response speed and error rate from the behavioral data by normalizing data prior to averaging. RTs were expressed as z scores, whereas error data were expressed as a difference from that subject's mean accuracy rate (z scores were not used for accuracy data, because accuracy is strongly nonnormal in distribution). We then calculated the mean RT and percentage of error for each category in both the behavioral and the simulated data. Model and behavioral data were compared across sequential categories through a 16-point correlation coefficient.

Conflict-control analysis. To determine the influence of conflict monitoring on control adjustment in the model and its relationship to behavioral performance, we compared the model with the behavioral data under two different simulation runs. The first was with the intact model, whereas the second was with the strategic priming mechanism (which modulates control as a function of conflict) lesioned. The value of strategic priming in the second run was fixed at the mean value over the first run. In order to reduce variability and provide a reliable measure of the predictive behavior of the model, each of these two conditions was run without noise. This had the effect of producing perfect accuracy in the model. Consequently, the maximum activity of the incorrect response representation was used as a measure of the predicted likelihood of an error. One problem with this approach is that subjects in experimental studies do make overt errors, presumably because behavioral performance is

an intrinsically noisy process. Conflict theory suggests that the noise present on error trials also produces increased conflict, even under exactly equivalent task conditions (owing to extra activity in the incorrect response channel; Botvinick et al., 2001). Consequently, removing this noise would change an important source of conflict and control adjustment. To compensate, we selectively sampled the conflict distribution of the model from a noisy run in order to determine the average conflict associated with errors under normal conditions. We then substituted this value into the model on trials in which the corresponding behavioral subject produced an error.

The ability of the intact model to account for additional variation in the trial-by-trial reaction time and error rates of individual subjects beyond that accounted for by the lesioned model was calculated as a partial correlation. The behavioral data used as targets for the simulation were taken from the equal-frequency 2AFC condition.

RESULTS

Frequency Effects

Simulating behavioral data. There was a strong main effect of frequency in the empirical data, in both the two-response and the one-response tasks. Poorer performance, in terms of slower RTs and greater errors (Figure 1), was observed for low-frequency stimuli, as compared with either high- or equal-frequency stimuli (all ps < .001). This
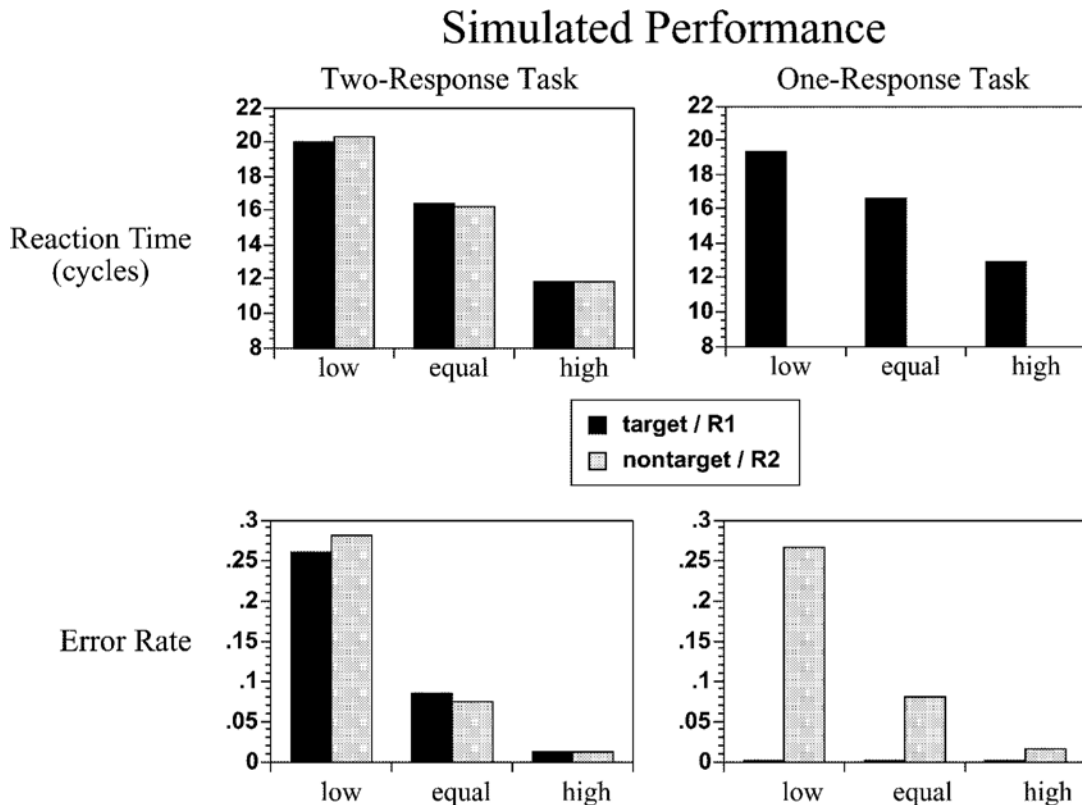

Figure 3. Performance of the model as a function of frequency (compare with Figure 1). The model captures the general trend toward poor performance as stimulus frequency decreases and also accounts for the exception present for target stimuli in the one-response conditions. Note that the target/R1 stimulus is low frequency when the nontarget/R2 stimulus is high frequency. Thus, the dark bar above the label "low" is taken from the same experiment as the light bar above the label "high," and vice versa. Both data points above the label "equal" are taken from the same experimental condition.

finding indicates that it is more difficult to respond appropriately to low-frequency events, which is consistent with the hypothesis that low-frequency events provide a source of conflict. The model was able to capture the pattern of behavioral data quite well in both qualitative and quantitative terms (Figure 3), displaying a parallel trend of performance decline for low-frequency events when compared against either high-frequency [RT, $F(1,63) = 3,771$, $p < .001$; accuracy, $F(1,63) = 2,563$, $p < .001$] or equal-frequency [RT, $F(1,63) = 413$, $p < .001$; accuracy, $F(1,63) = 1,430$, $p < .001$] events.

However, an exception to this frequency-dependent pattern of performance was observed in the empirical data for errors in the one-response low-frequency target task (i.e., the *oddball* task). In this condition, low-frequency responses were nearly perfectly accurate and were not significantly different from responses to high-frequency [$F(1,63) = 0.297$, $p = .588$] or equal-frequency [$F(1,63) = 0.133$, $p = .717$] events. Although, prior to simulations, this lack of a frequency effect appeared to be a potential problem for the conflict model, it was actually captured well [low vs. high, $F(1,63) = 3.07$, $p = .085$; low vs. equal, $F(1,63) = 1.72$, $p = .195$]. The model's behavior can be explained by the fact that activation of the incorrect (nontarget) decision layer unit does not lead to response execution. Thus, if the nontarget decision unit becomes transiently activated, it is not too late for the target decision unit to recover and lead to the execution of the correct response. This is in contrast to the two-response task, in which transient activation of the incorrect response leads to the production of an irreversible overt action.

Simulating neuroimaging data. A primary finding reported in Braver et al. (2001) was that there was significantly increased ACC activity for low-frequency stimuli, relative to high-frequency stimuli (see Figure 4). More-over, this pattern was observed in both the one and the two-response tasks. In contrast, under equal-frequency conditions, ACC activity did not differ across the two stimulus types (Figure 4). In the simulations of task performance, the conflict index from the model closely resembled the frequency-related pattern of ACC activation (see Figure 5). There was significantly greater conflict for low- versus high-frequency events in all task conditions (all ps $< .001$). Moreover, the conflict difference between low- and high-frequency stimuli was significantly greater than the difference between stimulus types under equal-frequency conditions [$F(1,63) = 9,143$, $p < .001$].

Sequential History

Simulating behavioral data. Our second analysis decomposed the behavioral data in terms of the specific five-trial stimulus sequence history. As was described above, previous studies have demonstrated that performance on two-response tasks can be significantly affected by subtle differences in sequential history. This effect can be observed in the present data as modulations of performance as a function of stimulus history (expressed as deviations from the mean, and for RT data additionally normalized into z-score units; see Figure 6). When the empirical data across one- and two-response tasks are compared, we see that the same general pattern holds. There are strong correlations between the two tasks in terms of the sequence–performance relationship (RT, $r = .933$; accuracy, $r = .908$). In Cho et al. (2002), we demonstrated that in two-response tasks, these performance modulations can be captured by the sequential priming mechanisms implemented in the model. Here, we see that the present model generalizes this property to the modified version of Cho et al. used here and to one-response tasks. Specifically, there is a high and roughly equivalent correlation
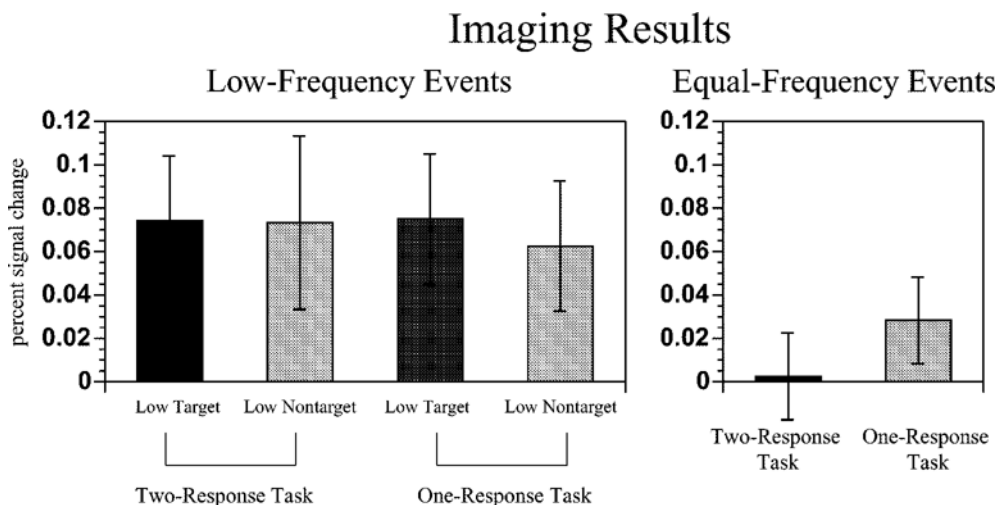

Figure 4. Neuroimaging results originally reported by Braver, Barch, Gray, Molfese, and Snyder (2001). The ACC shows a strong response for low-frequency events (low − high), but there is little difference in ACC activity between stimuli in the equal-frequency conditions.
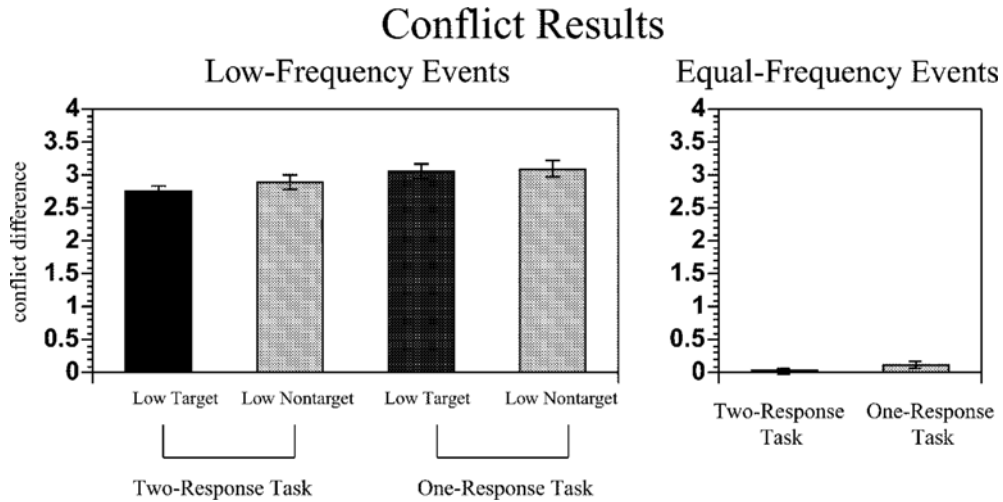
Conflict Results

Figure 5: Simulated conflict as a function of frequency (compare with Figure 4). Conflict is significantly higher for low-frequency stimuli (low – high), but there is no difference between stimuli in the equal-frequency conditions.

between model and empirical data for both the two-response (accuracy, $r = .890$; RT, $r = .933$; $ps < .001$) and the one-response (accuracy, $r = .942$; RT, $r = .897$; $ps < .001$) conditions.

The sequential history analysis makes clear the relationship between sequence type and performance. For example, the category RRRR, which corresponds to a trial on which the four previous stimuli match the stimulus on the current trial, had the fastest RT and the fewest errors. In fact, the four categories ending in RR had the four lowest RTs and error rates in both the behavioral and the simulated data. In contrast, the category RRRA, which corresponded to a trial on which the previous four stimuli were congruent, but the stimulus on the present trial mapped onto the alternate response (e.g., AAAAB), had the slowest RT and the most errors.

One question that arises from the model is whether the sequence–performance function is mirrored in the pattern of conflict across sequence categories. We examined this question by calculating the 16-point correlation between behavior in each sequence category and the corresponding conflict value for that category. The high correlation observed between conflict and empirical data behavioral performance (accuracy, $r = .934$; RT, $r = .891$) indicates that conflict not only tracks global frequency but also the local sequential history of stimuli. Interestingly, this result suggests a novel prediction of the model: that it is the local sequential history, rather than global frequency, that should drive modulations in ACC activity in both one-response and two-responses tasks. This prediction is testable by examining whether the equal-frequency conditions of the imaging data also showed modulations in ACC activity that appropriately tracked local sequence. It is important to note that this is a strong test of the model, given that no differences in ACC activity related to stimulus type were observed in the equal-frequency condition when

these data were analyzed at the global level (i.e., collapsing across stimulus histories).

Simulating neuroimaging data. To test this model prediction, we reanalyzed the imaging data by selectively averaging the equal-frequency condition data (including and collapsing across both two-response and one-response conditions to gain statistical power) according to the five-trial stimulus history. This analysis was complicated by our use of rapid event-related fMRI and the slow evolution of the hemodynamic response. This hemodynamic response lag creates overlap in the response from previous trials when these are closely spaced in time. Typically, the overlap can be disregarded when there is no systematic order in the sequential history of events (owing to counterbalancing). However, in the present case, sequential history was precisely the variable we wished to examine. Our approach for overcoming this limitation was to create difference timecourses from sequences that were matched in sequential history but differed in the final stimulus (e.g., RRRA–RRRR or AARA–AARR). We subtracted each of the eight sequence types ending in a repetition from the equivalent sequence ending in an alternation and determined the peak activation for each time course in each subject.

This procedure enabled quantification of the amplitude of ACC activity uniquely associated with alternation in each of eight possible sequential contexts. The prediction of the model was that alternation would yield significantly increased conflict primarily in contexts following a sequence of repetitions (i.e., RRRA–RRRR). This prediction is made clear in Figure 7, which plots the conflict data as eight paired differences in conflict between matched sequences. These difference scores demonstrate a coherent pattern of fluctuations in which conflict tends to decrease as the number of repetitions prior to an alternation decreases. However, it is also noteworthy that the pattern is
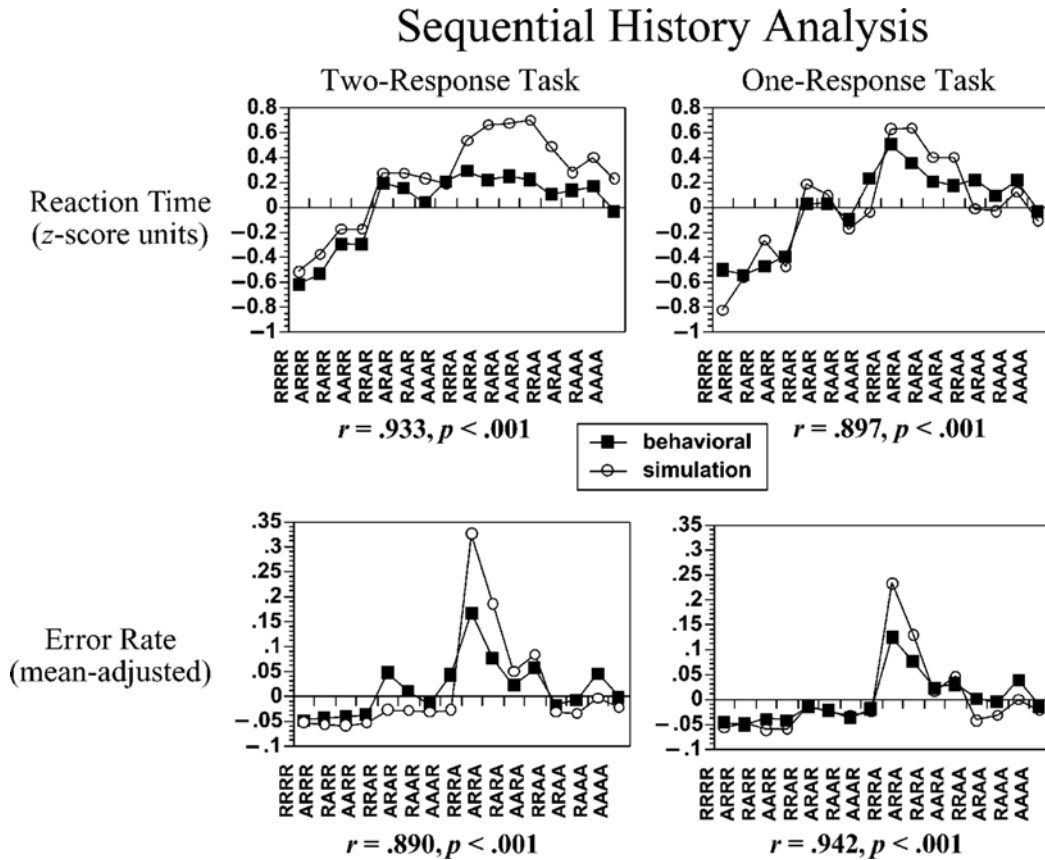
Figure 6: A more fine-grained comparison between simulation and behavioral performance. Each data point represents the mean for 1 of 16 sequential categories. RRRA, for instance, represents a sequence of three repetitions followed by one alternation.

not perfectly linear, demonstrating the complexity of the sequential history effects.

The results of the imaging data reanalysis are also shown in Figure 7, for comparison against the model. Although not a perfect match, the empirical data from the ACC corresponds rather well with the conflict predictions obtained from the model ($r = .856, p < .001$), accounting for 73% of the variance present in the sequential pattern of ACC activation.[1] It is noteworthy that the first two data points—those corresponding to sequences RRA–RRR—yielded the highest conflict in the model and the highest ACC activity. Recall that the RRA sequences are associated with poor behavioral and simulation performance (both slow reaction times and high error rates). In the model, this can be attributed to the high degree of priming present for the repeated stimulus that must be overcome when an alternation finally occurs. It is not surprising, then, that conflict is high for these trials. The fact that ACC activity shows a similar pattern of activation in this more fine-grained analysis demonstrates that the frequency-related ACC results reported in Braver et al. (2001) can indeed be attributed to local conflict dynamics and are not simply due to a gross effect of stimulus frequency. This finding is even more impressive when it is considered that Braver et al.'s study was not designed for the reanalysis that we performed on the data (and, consequently, had relatively low statistical power).

Conflict-Control Effects

Trial-by-trial predictions. In our third set of analyses, we attempted to obtain evidence that detection of conflict leads to control adjustments that influence behavioral performance on a trial-by-trial basis. To examine the role of control adjustments on the pattern of empirical performance data, we determined the ability of the conflict-control loop in the model to uniquely capture variance in the behavioral data. The average trial-by-trial correlation between the model and the subject data was .412 (99% confidence interval; .389–.435) for RT and was .254 (99% confidence interval; .225–.283) for errors (recall that the maximum activity in the incorrect response channel was used as a measure of the probability of error). Thus, on average, the model accounts for 16.6% of the variance present in the RT and 6.8% of the variance present in error
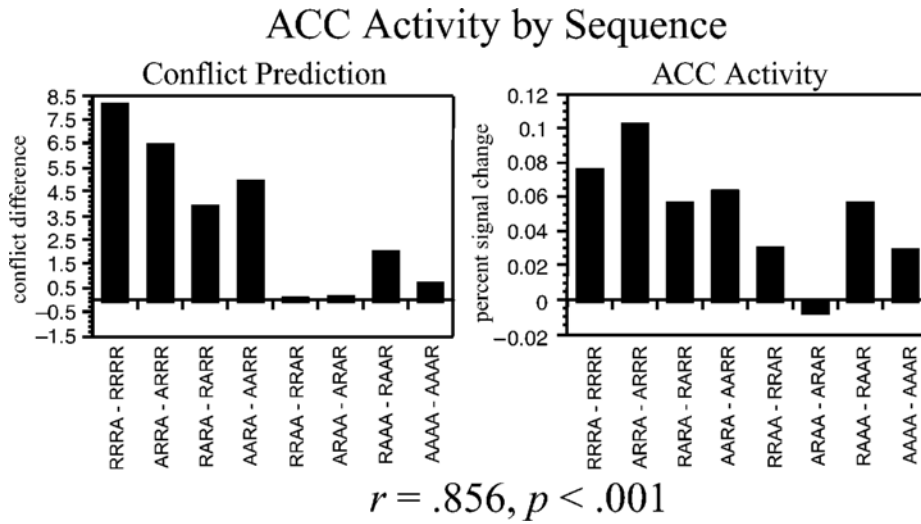
Figure 7. Predictions of the model with regard to conflict as a function of sequence and results of a reanalysis of the neuroimaging data. Each bar represents the difference between conflict or ACC activity for sequences matched on all but the final stimulus.

rate on any given trial for any given subject. Although these numbers indicate a somewhat low fit, it is important to note that noise processes in humans (which are poorly understood) may make a significant contribution to trial-by-trial behavioral performance. Although it is possible to present the simulation with the exact sequence of stimuli presented to a behavioral subject, it is not possible to simulate the exact pattern of neural or cognitive noise (which may actually be structured—i.e., non-Gaussian) experienced by that subject while performing the task. Thus, we expect that even a very good model of the underlying cognitive mechanisms involved in speeded response tasks

will capture only a small fraction of the total variance present in trial-by-trial performance.

The performance of the model with no conflict-monitoring control loop was slightly less effective at predicting performance, thus indicating a contribution of this component of the model. We calculated the specific contribution of the conflict-control loop by computing a partial correlation between the model and the behavioral data after controlling for the effect of the lesioned model. This partial correlation was computed separately for each subject and was positive for 64 out of 65 subjects in the RT analysis (mean partial correlation = .141; 99% confidence
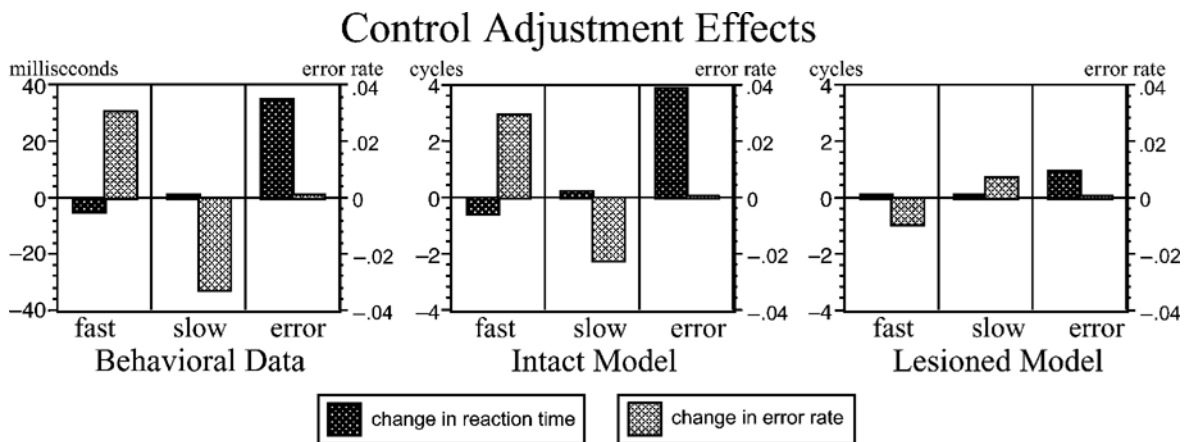
Figure 8. Control adjustments as a function of conflict. Changes in both reaction time and error rate performance are plotted in relation to three types of critical events: fast responses (low conflict), slow responses (high conflict), and errors (high conflict + error). Effects are shown for the behavioral data (left panel), the model with intact conflict-control loop (middle panel), and the model with lesioned conflict-control loop (right panel).

interval, .122–.160) and for 61 out of 65 subjects in the error analysis (mean partial correlation = .066; 99% confidence interval, .046–.087). Thus, although the conflict-control effect was rather small—accounting for 2.5% of the total variation in the subjects' RT scores and 1.5% of variance in error rates—it was highly reliable.

An important question arises as to what aspects of the behavioral data are captured more successfully by the control loop model. To gain insight into this issue, we attempted to operationally identify trials thought to be high versus low conflict and then to examine whether there were performance adjustments made in response to these events. On the basis of the previous literature on speeded response tasks (Laming, 1968; Rabbitt, 1966) and our own previous simulation studies (Botvinick et al., 2001; Yeung et al., 2002), we assumed that trials on which an error was made would have high conflict. However, we were also interested in exploring high-conflict trials in which no error was made. We postulated that trials on which RT was slower than the mean for that sequence type would have high conflict, whereas trials with fast RTs (again, relative to the sequence type mean) would have low conflict. Our hypothesis, implemented in the model, was that conflict, whether occurring in the presence of an error or not, would lead to a corresponding adjustment in control. This control adjustment would be manifest as a shift in the SATF: increased RT, but decreased error rates, following high-conflict events and conversely, decreased RT, but increased error rates, following low-conflict events.

We addressed this issue by examining performance in the model under intact and lesioned conditions and comparing this with the empirical data (Figure 8). We first binned trials according to sequence type and confirmed that, in the model, slow RT and error trials did have higher conflict than did fast RT trials. We then computed an index of control adjustment, by taking the difference in performance (both error rate and RT) between the trial immediately preceding and the trial immediately following the critical event for each of three categories (errors, slow RT, and fast RT). In the intact model, control adjustments were apparent in all three conditions. Following an error trial, there was a clear increase in RT and a slight decrease in error rate. Following a slow RT trial, there was a clear decrease in error rate and a slight slowing in RT. Conversely, following a fast RT trial, there was a clear *increase* in error rate, along with a slight decrease in RT. These effects were directly mediated by the presence of the conflict-control loop, since the lesioned model did not show such performance adjustments. Most importantly, the behavioral data showed a pattern of adjustment strikingly similar to that found in the intact model. There were significant adjustments in RT following errors [$t(60) = 6.50, p < .001$], in error rate following slow RT trials [$t(64) = -5.76, p < .001$], and in both RT and error rate following fast RT trials [RT, $t(64) = -2.6, p < .01$; error rate, $t(64) = 4.81, p < .001$]. The fact that the intact, but not the lesioned, model captured this pattern of behavioral performance suggests

that these effects are directly related to the conflict-control loop.

DISCUSSION

The primary motivation for this study was to provide increased support for our conflict theory of ACC function and, more generally, for the role that conflict plays in cognitive control. We were particularly interested in examining our theory within the context of simple speeded response tasks, which includes 2AFC, go/no-go and oddball as task variants. Our previous neuroimaging data suggested that the ACC was activated in an equivalent way across these three tasks by the occurrence of low-frequency stimuli (Braver et al., 2001). This finding was consistent with the idea that low-frequency events would be associated with increased conflict and that the ACC serves as a "generic" detector for the presence of such conflict. However, the equivalence of ACC activity across the three tasks also seemed somewhat puzzling from the standpoint of behavioral performance. In two-response tasks, subjects demonstrate a frequency-dependent decrease in accuracy, whereas no such modulation in accuracy is found in one-response tasks (at least for target responding). The discrepancy between the ACC activity, on the one hand, and the behavioral performance data, on the other, presented a challenge for the conflict theory. The key question we explored was whether a computational model implementing the conflict-monitoring hypothesis could simulate both the pattern of behavioral data and the pattern of conflict (as indexed by ACC activity in the imaging data) across one- and two-response tasks. To address these issues, we incorporated the mechanisms developed in Cho et al. (2002) into the conflict model described by Botvinick et al. (2001) and used these to simulate the effects of frequency on the target go/no-go, oddball, and 2AFC tasks.

The simulation results presented here suggest that our choice discrimination model is, in fact, able to capture data regarding both ACC activity (as an index of conflict) and behavioral performance in these tasks. The apparent discrepancy between the behavioral and the imaging data is resolved in the model by casting both tasks in terms of competing responses and subsequent conflict. More specifically, we have shown that the apparent inhibition involved in one-response tasks is also present in two-response tasks. In one-response tasks, a "no-go" representation must compete against and suppress the target response, just as the target response must compete against and suppress the alternative response pathway in two-response tasks. The difference in behavioral profiles across the two tasks is captured in the model by the simple fact that premature activation of the incorrect nontarget decision layer unit does not result in an overt behavior under a one-response task situation. The network can then recover from an error and still produce the correct response (something which is not possible in the two-response model, because by this

point the model has already generated an incorrect overt response). Thus, the model indicates that the two-response and one-response tasks are different at some level (as common sense suggests they have to be) but that this difference occurs downstream from where conflict is detected.

Sequential Effects on Conflict

In addition to capturing the relevant empirical data, the model also provides an explicit account of why ACC activity is increased to low-frequency stimuli. Specifically, in the model, repeated presentations of high-frequency stimuli create an expectancy for continued future repetitions, which results in residual activity in the expected stimulus–response pathway. The occurrence of the low-frequency stimulus must overcome the preactivated frequent response and, thus, produces conflict, reflected as increased ACC activity. Interestingly, however, the model suggests that these conflict effects can be accounted for by local sequential structure, rather than by global frequency. We observed that conflict fluctuated significantly even under conditions in which stimulus frequencies were globally equivalent (see Figure 7). Moreover, in the model, sequence-related fluctuations in conflict are mirrored in terms of behavioral performance. The ability of the model to capture the subtleties of sequential performance effects in the empirical data provides support for the behavioral component of the model. More important, this finding drove us to test the model's ability to predict a novel aspect of ACC activity that we had not previously considered prior to simulation work. In particular, the simulations suggested that ACC activity should fluctuate with sequential history just as behavioral performance did. We tested this prediction in a reanalysis of the imaging data, comparing ACC activity in the eight time course pairs matched for sequential history against the corresponding pattern of conflict derived from the model. This analysis yielded a correlation that was impressive ($r = .856$), especially given the post hoc nature of the analysis and its low statistical power. Critically, prior to simulations, this type of specific and quantitative prediction regarding the effect of sequential history on ACC activity would not have been possible. This is due to the complex nonlinear dynamics of the system. Thus, we were able to use the model not merely to replicate known phenomena, but also as a tool for generating testable new predictions about brain activity and behavior.

Given the close correspondence between the level of conflict in the model and the particular pattern of sequential history, one might wonder exactly what the conflict index that we hypothesize is computed in ACC is signaling during speeded response tasks. It is important to note that the forces that modulate the level of conflict in the ACC are the same ones that also modulate the level of sequential priming and, thus, behavioral performance. This explains the close correspondence between the conflict index for each sequential history category and behavioral performance in that category. Moreover, the conflict index is not the primary causal force that modulates behavioral performance. Instead, conflict merely reflects these fluctuations that are present elsewhere in the system in terms of changes in sequential priming mechanisms. Nevertheless, the local transformation of sequential history information (which is implicitly coded in the repetition and alternation priming mechanisms) into a conflict index within the ACC is very useful from a functional perspective. That is, by being actively represented within a neural region such as the ACC, conflict information can be used to directly modulate or bias the activity of other neural systems. In particular, we hypothesize that this conflict information is fed into systems responsible for adjusting control processes during task performance so as to optimize behavior.

Conflict and Cognitive Control Adjustments

A secondary goal of the present study was to directly test this hypothesis regarding the relationship between conflict and control adjustments in processing strategy. Specifically, we tested whether the model's success in accounting for trial-by-trial variations in behavioral performance was affected by the presence or absence of a control module that regulates strategic priming on the basis of input from a conflict detector. We clearly determined, in two separate analyses, that the presence of the conflict-driven control module did contribute to the success with which the model accounted for behavioral performance. First, we demonstrated that in the individual-subject data, there was a small, but highly statistically reliable, increase in variance accounted for by the model with a control loop, as compared with a model without this loop (improvements found in 64/65 subjects for RT and 61/65 for errors). Second, in the group-averaged behavioral data, we demonstrated the presence of highly significant control adjustment effects (see Figure 8) that could be captured only when the conflict-control loop was present in the model.

Together, these results provide compelling evidence that sequential modulations in conflict lead to adjustments in control strategies that are reflected in behavioral performance. In particular, the simulations suggest that conflict information can serve as a useful index for optimizing performance, by indicating when there is either too much or too little strategic priming, given the present sequential structure of events. This ability to adjust performance in an on-line manner in accordance with situational contingencies is a hallmark of cognitive control. Our model provides a tool for revealing the subtle presence of cognitive control processes at work during even simple speeded response tasks. Critically, the control adjustments observed in the present study go beyond previous work on performance adjustments in speeded response tasks (Laming, 1968; Rabbitt, 1966), by demonstrating that these adjustments can occur even when performance is error free but conflict is high nonetheless. Furthermore, the finding of conflict-related sequential relationships in the behavioral data also makes clear that there is an underlying autocorrelative component to performance in these tasks. It is interesting that recent attention has been drawn to the ubiquitous presence and significance of autocorrelation within human performance data (D. L. Gilden, Thornton,

& Mallon, 1995; F. Gilden, 2001). Out of this work, it has been suggested that the particular patterns of behavioral autocorrelation reflect the presence of internal representational structure and nonlinear dynamics within the cognitive system. Our model provides just such an explicit representational mechanism by linking sequential relationships in performance to a feedback loop between conflict detection and cognitive control adjustment.

The ability of the control loop model to predict extra variance in the trial-by-trial performance of individual subjects provides a novel method of evaluating individual differences in control function. In particular, these estimates can be conducted for subjects in different populations suspected of showing dysfunction in cognitive control (e.g., patients with schizophrenia, older adults, etc.). Such dysfunction would be expected to show up in terms of the goodness of fit of the model with a control-loop versus the lesioned model (determined through the partial correlation procedure described above). Importantly, the use of the model would provide a rigorous means for detecting behavioral disturbances that might be quite subtle in such tasks (given that the control loop accounts only for about 2% of variability in overall performance). More important, the use of the model in this manner would provide a means for linking together theories of cognitive dysfunction in these populations to biologically based theories relating conflict monitoring and control adjustment to specific neural and computational mechanisms.

Neural Mechanisms

Despite the *computational* specificity of our model, up to this point we have said very little about the biological structures underlying its components, with the exception of the ACC. However, some of the assumptions we have made with regard to functional components of the model have important implications regarding neural substrates. For example, we have argued that there exists a functional component involved in response selection that actively represents both *go* and *no-go* action representations. The assumption that no-go actions are actively represented is somewhat counterintuitive, because the same computation could be performed through attenuation or inhibition of activity in a *go* representation. However, there is some support for this assumption from neurophysiology. In particular, Hanes, Patterson, and Schall (1998) have recently demonstrated that neurons within the supplementary motor cortex of macaque monkeys exhibit activity only during the successful suppression of a partially prepared response. Thus, it may be that the response selection and suppression functions implemented in our model are carried out within higher order motor regions of the cortex. Moreover, the conceptualization of processing in go/no-go tasks being one of competition between two active representations, rather than suppression of a single representation, is consistent with the emerging general framework of neural processing and representational competition that has been discussed in recent theoretical treatments (Desimone & Duncan, 1995; Miller & Cohen, 2001).

We have also argued for the presence of a strategic priming component that provides a nonspecific preparatory input to the response selection layers modulated by the degree of previous conflict experienced. We chose to model speed–accuracy adjustments in this way to maintain consistency with former models (Botvinick et al., 2001) and with neurophysiological data that indicate that RTs are related to the rate of accumulation of neural information, rather than to variations in the threshold of response (Hanes & Schall, 1996). Nevertheless, a more thorough exploration of the impact and neurobiological plausibility of various strategic control mechanisms is a fruitful topic for further research.

The strategic priming functions of the model are similar to the constructs of sustained attention and arousal. Given the long-standing hypothesis that sustained attention and arousal are subserved by the right-lateralized prefrontal and parietal cortex (Posner & Petersen, 1990), these regions would serve as natural candidates for the locus of the strategic priming signal. This hypothesis is also consistent with our recent neuroimaging data suggesting that the right-hemisphere dorsolateral prefrontal cortex and the inferior parietal cortex are engaged, along with the ACC, by low- frequency events during speeded response tasks (Braver et al., 2001). It is also possible that neuromodulatory systems, such as the brainstem noradrenergic nucleus (locus coeruleus), might be responsive to conflict detection (Aston-Jones, Rajkowski, & Cohen, 2000; Cohen, Botvinick, & Carter, 2000) and, thus, be in a position to regulate decisional and motor processes (Usher, Cohen, Servan-Schreiber, & Rajkowski, 1999).

CONCLUSION

The present study was not intended to provide a complete account of the myriad processes that contribute to cognitive control. However, it does make significant strides toward elucidating a specific type of control that shows up in simple cognitive tasks. Furthermore, by separating control itself from indexing the need for control through conflict detection, we have provided an explicit model that makes no reference to a homunculus. By using the model to account for detailed aspects of behavioral and neuroimaging data, we have drawn a tight link between the functional and dynamical properties of conflict detection and the activation of a specific brain region, the ACC. Finally, the utility of the model reaches beyond the scope of this study. The present work suggests new predictions regarding ACC function and activity, provides a potential diagnostic tool for assessing control function in clinical populations, and constrains the development of theories regarding the neural substrates of control.

REFERENCES

Aston-Jones, G., Rajkowski, J., & Cohen, J. (2000). Locus coeruleus and regulation of behavioral flexibility and attention. *Progress in Brain Research*, **126**, 165-182.

Barch, D. M., Braver, T. S., Akbudak, E., Conturo, T., Ollinger, J., &

SNYDER, A. V. (2001). Anterior cingulate cortex and response conflict: Effects of response modality and processing domain. *Cerebral Cortex*, **11**, 837-848.

BARCH, D. M., BRAVER, T. S., SABB, F. W., & NOLL, D. C. (2000). The anterior cingulate cortex and response competition: Evidence from an fMRI study of overt verb generation. *Journal of Cognitive Neuroscience*, **12**, 298-305.

BOTVINICK, M. M., BRAVER, T. S., BARCH, D. M., CARTER, C. S., & COHEN, J. C. (2001). Conflict monitoring and cognitive control. *Psychological Review*, **108**, 624-652.

BOTVINICK, M. M., NYSTROM, L., FISSEL, K., CARTER, C. S., & COHEN, J. D. (1999). Conflict monitoring versus selection-for-action in anterior cingulate cortex. *Nature*, **402**, 179-181.

BRAVER, T. S., BARCH, D. M., GRAY, J. R., MOLFESE, D. L., & SNYDER, A. (2001). Anterior cingulate cortex and response conflict: Effects of frequency, inhibition, and errors. *Cerebral Cortex*, **11**, 825-836.

BUSH, G., LUU, P., & POSNER, M. I. (2000). Cognitive and emotional influences in anterior cingulate cortex. *Trends in Cognitive Sciences*, **4**, 215-222.

CARTER, C. S., BOTVINICK, M. M., & COHEN, J. D. (1999). The contribution of the anterior cingulate cortex to executive processes in cognition. *Reviews in the Neurosciences*, **10**, 49-57.

CARTER, C. S., BRAVER, T. S., BARCH, D. M., BOTVINICK, M. M., NOLL, D. C., & COHEN, J. D. (1998). Anterior cingulate cortex, error detection, and the online monitoring of performance. *Science*, **280**, 747-749.

CARTER, C. S., MACDONALD, A. M., BOTVINICK, M., ROSS, L. L., STENGER, A., NOLL, D., & COHEN, J. D. (2000). Parsing executive processes: Strategic versus evaluative functions of the anterior cingulate cortex. *Proceedings of the National Academy of Sciences*, **97**, 1944-1948.

CASEY, B. J., THOMAS, K. M., WELSH, T. F., BADGAIYAN, R., ECCARD, C. H., JENNINGS, J. R., & CRONE, E. A. (2000). Dissociation of response conflict, attentional selection, and expectancy with functional magnetic resonance imaging (fMRI). *Proceedings of the National Academy of Sciences*, **97**, 8728-8733.

CASEY, B. J., TRAINOR, R. J., ORENDI, J. L., SCHUBERT, A. B., NYSTROM, L. E., GIEDD, J. N., CASTELANOS, F. X., HAXBY, J. V., NOLL, D. C., COHEN, J. D., FORMAN, S. D., DAHL, R. E., & RAPOPORT, J. L. (1997). A developmental functional MRI study of prefrontal activation during performance of a go–no-go task. *Journal of Cognitive Neuroscience*, **9**, 835-847.

CHO, R. Y., NYSTROM, L. E., JONES, A. D., BRAVER, T. S., HOLMES, P. J., & COHEN, J. D. (2002). Mechanisms underlying dependencies of performance on stimulus history in a two-alternative forced-choice task. *Cognitive, Affective, & Behavioral Neuroscience*, **2**, 283-299.

COHEN, J. D., BOTVINICK, M., & CARTER, C. S. (2000). Anterior cingulate and prefrontal cortex: Who's in control? *Nature Neuroscience*, **3**, 421-423.

DESIMONE, R., & DUNCAN, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience*, **18**, 193-222.

GILDEN, D. L., THORNTON, T., & MALLON, M. W. (1995). 1/f noise in human cognition. *Science*, **267**, 1837-1839.

GILDEN, F. (2001). Cognitive emissions of 1/f noise. *Psychological Review*, **108**, 35-56.

HANES, D. P., PATTERSON, W. F., II, & SCHALL, J. D. (1998). Role of frontal eye fields in countermanding saccades: Visual, movement, and fixation activity. *Journal of Neurophysiology*, **79**, 817-834.

HANES, D. P., & SCHALL, J. D. (1996). Neural control of voluntary movement initiation. *Science*, **274**, 427-430.

HOPFIELD, J. J. (1982). Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the National Academy of Sciences*, **79**, 2554-2558.

KIRBY, N. (1980). Sequential effects in choice reaction time. In A. T. Welford (Ed.), *Reaction times* (pp. 129-172). London: Academic Press.

LAMING, D. R. J. (1968). *Information theory of choice reaction times.* London: Academic Press.

LUCE, R. D. (1986). *Response times: Their role in inferring elementary mental operations.* New York: Oxford University Press.

MACDONALD, A. W., COHEN, J. D., STENGER, V. A., & CARTER, C. S. (2000). Dissociating the role of the dorsolateral prefrontal cortex and anterior cingulate cortex in cognitive control. *Science*, **288**, 1835-1838.

MCCLELLAND, J. L. (1979). On the time relations of mental processes: An examination of systems of processes in cascade. *Psychological Review*, **86**, 287-330.

MILLER, E. K., & COHEN, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience*, **21**, 167-202.

PAUS, T., KOSKI, L., CARAMANOS, Z., & WESTBURY, C. (1998). Regional differences in the effects of task difficulty and motor output on blood flow response in the human anterior cingulate cortex: A review of 107 PET activation studies. *NeuroReport*, **9**, R37-R47.

POSNER, M. I., & DIGIROLAMO, G. J. (1998). Executive attention: Conflict, target detection and cognitive control. In R. Parasuraman (Ed.), *The attentive brain* (pp. 401-423). Cambridge, MA: MIT Press.

POSNER, M. I., & PETERSEN, S. E. (1990). The attention system of the human brain. *Annual Review of Neuroscience*, **13**, 25-42.

RABBITT, P. M. A. (1966). Errors and error correction in choice-response tasks. *Journal of Experimental Psychology*, **71**, 264-272.

RABBITT, P. M. A., & VYAS, S. M. (1970). An elementary preliminary taxonomy for some errors in laboratory choice RT tasks. *Acta Psychologica*, **33**, 55-76.

REMINGTON, R. J. (1969). Analysis of sequential effects in choice-reaction times. *Journal of Experimental Psychology*, **82**, 250-257.

SMITH, G. A., & BREWER, N. (1995). Slowness and age: Speed–accuracy mechanisms. *Psychology & Aging*, **10**, 238-247.

SOETENS, E., BOER, L. C., & HUETING, J. E. (1985). Expectancy or automatic facilitation? Separating sequential effects in two-choice reaction time. *Journal of Experimental Psychology*, **11**, 598-616.

USHER, M., COHEN, J. D., SERVAN-SCHREIBER, D., & RAJKOWSKI, J. (1999). The role of locus coeruleus in the regulation of cognitive performance. *Science*, **283**, 549-554.

USHER, M., & MCCLELLAND, J. L. (2001). The time course of perceptual choice: The leaky, competing accumulator model. *Psychological Review*, **108**, 550-592.

YEUNG, N., BOTVINICK, M. M., & COHEN, J. D. (2002). *The neural basis of error detection: Conflict monitoring and the error-related negativity.* Manuscript submitted for publication.

NOTE

1. The high correlation between the simulated conflict measure and ACC activity levels across sequential history conditions provided an opportunity to test how well alternative measures of conflict might capture the same data. In particular, besides the Hopfield energy measure that we used [i.e., $E = \sum_n X(n) \cdot Y(n)$], other conflict measures are also possible, such as (1) $E = \sum_n 1 - |X(n) - Y(n)|/[X(n) + Y(n)]$ and (2) $E = \sum_n [X(n) + Y(n)]^2/\{[X(n) - Y(n)]^2 + 1\}$. We implemented these alternative measures of conflict in the model and resimulated the sequential history data. When comparing the new conflict values against the ACC activity patterns, we observed correlations of $r = .859$ for the first tested measure and $r = .878$ for the second tested measure. The equivalent success of each measure in accounting for ACC data suggests that at least for the present stage of model-building, the exact form of the conflict computation is not critical.

APPENDIX

Below are the equations that govern activation dynamics, conflict, control adjustment, and sequential priming. In all of the equations, t represents the index of the current trial, and n represents the index of the present cycle within a trial. Table A1 gives a summary of model parameters and their values and highlights which parameters were held constant from previous simulations, which were changed in order to maintain consistency with previous models, and which were manipulated in preliminary simulations in order to match the behavioral data.

Activation Dynamics

Activity of the competing units in the decision layer and response execution layer were governed by leaky integrator and lateral inhibitory dynamics, which have been well studied in the theoretical literature (reviewed in Usher & McClelland, 2001). Unit activation was updated on a cycle-by-cycle basis according to the following equation:

$$X_t(n) = X_t(n-1)$$
$$+ \tau \cdot \left(-\left[\kappa \cdot X_t(n-1) \right] - \left\{ \beta \cdot f\left[Y_t(n-1) \right] \right\} + \zeta_t(n) + I_t(n) + S_t + B_t^X \right),$$

where $X_t(n)$ refers to the activity of the unit at cycle n, $Y_t(n)$ refers to the activity of the other competing unit in the layer, and $f(x)$ is a linear activation function. τ is a time constant, κ is a standard leak parameter, β is the strength of lateral inhibition, $I_t(n)$ is excitatory input from the preceding layer, and $\zeta_t(n)$ is a noise parameter with zero-mean and standard deviation σ. S_t and B_t are priming terms that changed from trial to trial according to conflict and sequence dynamics and will be described further below. The threshold value for the response execution layer was θ.

For units in the input layer, $X_t(n)$ and $Y_t(n)$ were clamped at fixed values (0–1 range) representing the presence or absence of an external stimulus. During the preparatory period (20 cycles), these values were set at 0. During the stimulus presentation period (30 cycles), the input unit representing the presented stimulus received a value of ρ; the other input received a value of $1-\rho$. These values were taken from Botvinick et al. (2001) and reflect a level of stimulus ambiguity associated with perceptual processing (furthermore, they provide a means of introducing variability owing to noise). Following stimulus presentation, activity was allowed to settle for an additional 70 cycles.

Our use of an additional processing layer was a significant alteration of the basic Usher et al. (2001) model, and we found it necessary to change some of the model parameters in order to maintain the basic behavior of the model (see Table A1). Specifically, we found it necessary to use a larger value of the time constant τ for the response execution layer. This increased the rate at which information was integrated at the response execution layer and corrected for sluggish behavior during the preparatory phase. We also decreased the strength of lateral inhibition, β. Because inhibition acts at both the response decision and the response execution layers, their combined effect caused activity in the execution layer to become suppressed more quickly than in the one-layer

Table A1
Model Parameters

Parameter	Description	Value	Status
	Basic Parameters		
τ	time constant (decision layer)	0.1	F
τ	time constant (execution layer)	0.2	C
β	inhibition	0.6	C
κ	leak parameter	0.25	F
ρ	stimulus ambiguity	0.85	F
σ	standard of noise	0.23	M
θ	response threshold	2.3	M
b	baseline "no-go" unit input	0.2	A
	Sequential Priming Parameters		
γ	time constant	0.5	F
M_R	maximum repetition priming	0.06	M
M_A	maximum alternation priming	0.02	M
c	number of priming cycles	35	A
	Strategic Priming Parameters		
λ	time constant	0.75	F
α	slope	-0.05	F
μ	constant intercept	0.5	F

Note—The status column indicates whether the parameter was fixed at a value taken from previous models (F), altered in order to mainhtain consistency with previous models (C), assigned a value in an appropriate range (A), or adjusted during preliminary simulations to match behavioral data (M).

model. By reducing this parameter, we were able to obtain activation trajectories at the execution layer that more closely resembled those of the one-layer model.

In the one-response version of the model, the inactive response unit was given a small constant baseline input ($b = .2$) so that it could compete with the active response unit. This parameter was not adjusted to match specific data but was chosen so that competition was strong enough to produce errors in the active response unit without the competing unit's ever crossing threshold.

Sequential Priming

Priming based on stimulus history was governed by the B_t^x parameter. The implementation that we used was termed IR1–SA2 by Cho et al. (2002). IR1 refers to "Individual Repetition One," a scheme in which each individual stimulus has its own repetition detector. The "one" refers to the fact that only a single presentation of Stimulus X is required in order to increase priming for X, as opposed to a scheme in which an actual repetition XX would have to occur before priming increased for X (this latter scheme was referred to as IR2). SA2 refers to "Shared Alternation Two," in which there exists a single alternation detector, which applies priming to either Stimulus X or Y, depending on which stimulus would be an alternation of the most recent stimulus. In order for the amount of alternation priming to increase, an actual alternation (XY or YX) had to be present in the recent history (hence, the "two"). For both repetition and alternation priming, the lack of a detected repetition or alternation caused the amount of priming to decay toward zero. These priming mechanisms are represented by the following equations:

$$R_t^X = \gamma \cdot R_{(t-1)} X + (1 - \gamma) \cdot M_R \cdot \delta_{(t-1)}^X$$

and

$$A_t = \gamma \cdot A_{(t-1)} + (1 - \gamma) \cdot M_A \cdot \varepsilon_{(t-1)},$$

where R_t^x refers to the repetition priming for Response X on trial t, A_t refers to the amount of alternation priming on trial t. Note the superscript for repetition priming because there were separate repetition priming mechanisms for each stimulus. $\gamma = 0.5$ is a time constant of integration. M_R and M_A are the maximum values of repetition and alternation priming respectively. δ_t^x and ε_t are binary step functions that served as the specific repetition and alternation detectors. δ_t^x took a value of 1 if the response on trial t was X and 0 if the response on trial t was not X. ε_t took a value of 1 if the response on trial t was an alternation of the response on trial $t-1$ and 0 if the response on trial t was not an alternation. The values of M_R and M_A were adjusted during preliminary simulations in order to fit the empirical data (see below). Repetition and alternation priming were simply added to produce the B_t^x parameter in the activation equation. That is, when X was the same stimulus as that presented on the previous trial, B_t^x was equal to R_t^x; when X was an alternation of the previous stimulus, B_t^x was equal to $R_t^x + A_t$.

Because of the linear activation function used in the model, allowing priming to endure for a full 120 cycles caused activity in the decision and execution layers to increase well beyond threshold levels. This created artificially high levels of conflict near the end of individual trials. To correct for this, we introduced an additional parameter, $c = 35$, the number of priming cycles. Thus, priming began at the start of the trial (during the preparatory phase) and lasted a total of 35 cycles. The value of this particular parameter is an artifact of our specific implementation (i.e., the choice of a linear activation function), and we feel that it has little relevance to actual psychological or biological phenomena.

Conflict

Conflict on trial t was measured by the following equation:

$$E_t = \sum_n X(n) \cdot Y(n),$$

where n is the cycle index and $X(n)$ and $Y(n)$ are the activity of the two decision layer units on cycle n. Thus conflict represents the joint product of activity in the two response decision layer units summed over all 120 cycles.

Strategic Priming

Conflict affected strategic priming as an exponentially weighed average according to the following equation taken from Botvinick et al. (2001):

$$S_t = \lambda \cdot S_{(t-1)} + (1 - \lambda) \cdot \left[\alpha E_{(t-1)} + \mu \right],$$

where S_t refers to the strategic priming activity on trial t and E_t refers to the conflict (energy) on trial t. The exponentially weighted averaging had the effect that conflict at time t has the largest effect on strategic priming at time $t + 1$, but it also continued to influence priming for a number of subsequent trials. The λ is a time constant controlling the rate at which the influence of conflict on previous trials decays. The α and μ are the scaling parameter and the constant intercept, respectively. If $E_{(t-1)}$ were zero, the value of S_t would tend toward μ; for higher values of $E_{(t-1)}$, S_t tends toward $[\alpha E_{(t-1)} + \mu]$. Since $\alpha < 0$, this implies that higher conflict on trial $t-1$ is associated with lower values of strategic priming on trial t.

APPENDIX (Continued)

Preliminary Simulations

We performed three preliminary simulations prior to those discussed in the results. In the first simulation, we looked at baseline error rates in the equal probability conditions and adjusted the response threshold θ and the standard deviation of noise σ until the model could capture this baseline error rate. Next, we were interested in the RT and error performance of the model as a function of stimulus frequency (Figure 1). To improve this fit, we adjusted the strength of repetition priming, M_R. Alternation priming had little effect on performance as a function of frequency but did influence the results of the sequential history analysis. In order to select an appropriate value of M_A, we used a grid search optimization procedure to maximize the fit of the model to the RT data as a function of frequency. Thus, four parameters in all were adjusted to fit the behavioral data, and these adjustments were made without regard for the ability of the model to predict activity in the ACC through the conflict measure.