Chapter 3

Spiking Neuron Simulations of the Relationships Among Repetition Suppression, Neural Synchronization, and Behavioral Priming

The following series of five simulations will develop the idea that short-term plasticity processes of firing-rate adaptation and synaptic depression are responsible for short-term repetition suppression effects in monkeys and humans. The explanation of how repetition suppression phenomena are related to behavioral priming effects hinges on the synchronizing role of short-term plasticity mechanisms in cortical networks, simultaneously affording better behavioral performance at a lower metabolic cost. The basic approach of the simulations is to examine the impact of repeatedly presenting a single "visual stimulus" to an artificial neural network of coupled excitatory and inhibitory spiking cells that is intended to capture basic aspects of the neural processing and representation of that stimulus. Changes in firing rate and spike synchronization across the network will be measured, and the importance of various neural processes (e.g. inhibition, adaptation, synaptic depression) for these results will be examined. These
dynamic activity changes will then be related to a simple measure of reaction time, and the degree of processing efficiency in the model across repetitions will be assessed.

3.1 General Simulation Methods: Simulations 1-5

Large networks of single-compartment, integrate-and-fire (IAF) spiking neurons were used for these simulations. One thousand spiking excitatory cells served as inputs and projected probabilistically [p(connect)=0.05] to 250 excitatory and 50 inhibitory cells, which were sparsely interconnected [p(connect)=0.3] such that 85% of connections and cells were excitatory and 15% were inhibitory, consistent with anatomical investigations of neocortex (e.g. Abeles, 1991; White, 1989). While this connectivity is slightly more dense than observed in actual neocortex [p(connect) ≈ 0.10], larger networks with more appropriate connection densities have yielded comparable results (results not shown), and the current density with a smaller network was chosen to afford more rapid simulation run times. The density of connections within and between the excitatory and inhibitory groups was also altered slightly to produce firing rates within a desired range of 30-40 spikes/second (Hz), given fixed constraints on the characteristics of the synaptic currents (discussed below) and the overall ratio of the number of excitatory to inhibitory connections (approximately 5/1). The architecture of the network is depicted graphically in Figure 3.1.
Figure 3.1: Architecture of the basic model used for Simulations 1-5. Poisson-spiking excitatory input neurons (N=1000) provide input to 250 excitatory and 50 inhibitory spiking cells that are sparsely interconnected.

The low connection probability of each input cell to the excitatory and inhibitory cells was chosen specifically to minimize shared inputs to individual cells in the coupled pool. This reduced the chance that any synchronous firing in the coupled pool was due simply to common inputs and insured that changes in synchrony were due to the dynamical interactions of cells in the coupled pool. However, this does not rule out the possibility that common inputs serve as one source of synchronous firing in the real brain; it simply provides an argument for the role of dynamical interactions in synchronous firing. In order to investigate the impact of heterogeneity in connectivity on repetition-related changes in synchrony, the number of incoming excitatory connections for cells in the coupled pool was varied randomly from cell to cell for different levels of heterogeneity.
four levels in all. The amount of input heterogeneity ranged from completely homogeneous (all cells received the same number of excitatory connections) to very heterogeneous by adding or subtracting a certain number of excitatory connections that was drawn randomly from a uniform distribution for each cell (4 levels: +/- 0%, 20%, 40%, or 60% of the number of excitatory inputs in the homogeneous case).

The parameters of the excitatory and inhibitory IAF neurons were set to be consistent with a wide range of empirical results from *in vitro* and (where possible) *in vivo* neurophysiological experiments of neocortical neurons. The passive membrane properties of the excitatory cells were matched to data on neurons classified physiologically as Regular Spiking Units (RSUs), believed to include pyramidal cells and spiny stellate cells, and the inhibitory cells were matched to data on interneurons classified as Fast Spiking Units (McCormick, Connors, Lighthall, & Prince, 1985):

e's: \[ C_m \frac{dV}{dt} = -g_{\text{leakE}} \cdot (V - V_{\text{Leak}}) + I_{\text{K(Ca)}} + I_{\text{AMPA}} + I_{\text{GABAa}} + I_0 \]
i's: \[ C_m \frac{dV}{dt} = -g_{\text{leakI}} \cdot (V - V_{\text{Leak}}) + I_{\text{AMPA}} + I_{\text{GABAa}} + I_0 \]

where \( C_m \) is the specific membrane capacitance (1 \( \mu \)F/cm\(^2\)), \( V \) is the instantaneous membrane potential (in mV), \( g_{\text{leakE}} \) (0.05 mS/cm\(^2\)) and \( g_{\text{leakI}} \) (0.1 mS/cm\(^2\)) are the constant leak conductances for excitatory and inhibitory neurons (set to be consistent with membrane time constants, \( \tau_m \), of 20 ms and 10 ms, respectively), \( V_{\text{Leak}} \) is the reversal potential of the leak current (-70 mV) that is responsible for the resting potentials of all the cells, and \( I_0 \) is a constant current representing the background activity from other cortical cells that placed the membrane potentials near the spiking threshold of -54 mV.
(e's: 0.8 µA/cm², i's: 1.6 µA/cm²). When \( V \) exceeded the spike threshold due to fluctuations in the synaptic currents coming from other cells in the network (\( I_{\text{AMPA}} \) and \( I_{\text{GABA}_a} \)), a spike was considered to have occurred, and \( V \) was instantaneously set to \( V_{\text{reset}} \) (e's: -60 mV, i's: -62 mV) for the duration of the cell's refractory period (e's: 2 ms, i's: 1 ms). These values were determined using the method of Troyer and Miller (1997) to generate spike-frequency versus current curves (\( f-I \)) for excitatory and inhibitory IAF neurons that matched those determined empirically for RSU and FSU neocortical cells (McCormick et al., 1985). The calcium-dependent potassium current (\( I_{K(Ca)} \)) responsible for firing-rate adaptation in excitatory cells and the fast excitatory (\( I_{\text{AMPA}} \)) and inhibitory (\( I_{\text{GABA}_a} \)) synaptic currents for all cells are discussed in detail below. The model was implemented in the C programming language, and all differential equations were integrated using the 2nd-order Runga-Kutta algorithm with linear interpolation to find precise spike times; the discretization of time was small enough to ensure accurate synchronization results (\( dt=0.02 \) ms; Hansel et al., 1998; Press, Teukolsky, Vetterling, & Flannery, 1992; Shelley & Tao, 2001).

### 3.1.1 Synaptic Currents

The fast excitatory (\( I_{\text{AMPA}} \)) and inhibitory (\( I_{\text{GABA}_a} \)) synaptic currents for each cell are given by:

\[
I_{\text{AMPA}} \, i = - \sum_j g_{ij} \cdot s_{\text{AMPA}, j} \cdot D_j \cdot (V - V_{\text{AMPA}})
\]

\[
I_{\text{GABA}_a} \, i = - \sum_k g_{ik} \cdot s_{\text{GABA}_a, k} \cdot D_k \cdot (V - V_{\text{GABA}_a})
\]
where \( g_{ij} \) is the maximal synaptic conductance from neuron \( j \) to neuron \( i \) (\( g_{ee}=0.02, g_{se}=0.025, g_{se}=0.15, g_{ii}=0.1, \) all in \( \text{mS/cm}^2 \)), \( s_{\text{AMPA}j} \) and \( s_{\text{GABA}k} \) are the dynamic \( s \) gating variables for the synaptic currents coming from the \( j \)-th excitatory and the \( k \)-th inhibitory neurons (discussed below), \( D_j \) and \( D_k \) are the dynamic synaptic depression scaling terms (ranging from 1 down to 0) from the \( j \)-th excitatory and \( k \)-th inhibitory neurons (discussed below in Section 3.1.2) and \( V_{\text{AMPA}} \) and \( V_{\text{GABA}} \) are the reversal potentials for the synaptic currents (0 and -80 mV, respectively). As indicated, the total AMPA and GABA currents for neuron \( i \) were calculated by summing across the individual currents due to each sending neuron. The maximal synaptic conductances were parameterized to produce peak excitatory postsynaptic potentials (EPSPs) of approximately 1 mV (from rest) and peak inhibitory postsynaptic potentials (IPSPs) of approximately 1-2 mV (from spike threshold) (see also Troyer & Miller, 1997). The dynamics of each \( s \) gating variable were governed by a system of two first-order differential equations that allowed the detailed time course of the gating variables to be matched with physiological data (see Wang & Rinzel, 1992; Wang, 1999, for discussion):

\[
\frac{dx}{dt} = \Sigma_i \delta(t - t_i^s) - x/\tau_x
\]

\[
\frac{ds}{dt} = \alpha_s \cdot x \cdot (1 - s) - s/\tau_s
\]

where \( \delta(t - t_i^s) \) is the Dirac \( \delta \) function used to increment \( x \) instantaneously by 1 at the time of each spike \( (t_i^s) \), \( \tau_x \) is the time constant for the decay of \( x \) back to 0 which also gives the approximately exponential rise time of \( s \) (\( \tau_{\text{AMPA}}=0.33 \text{ ms}, \tau_{\text{GABA}}=1 \text{ ms} \)), \( \alpha_s \) is a parameter that scales the height of \( s \) without substantially changing its time course.
(α_{\text{AMPA}} = 1.22, \, \alpha_{\text{GABAa}} = 0.152), and \, \tau_i \text{ is the time constant for the exponential decay of } s \text{ back to } 0 (\tau_{\text{AMPA}} = 3 \, \text{ms}, \, \tau_{\text{GABAa}} = 7 \, \text{ms}). The time constants for the rise and decay of AMPA and GABAa were taken from intracellular recording studies in cortex (e.g. Hausser, Major, & Stuart, 2001; Lumer, Edelman, & Tononi, 1997). The term of (1-s) in the equation for ds/dt insured that s would saturate below 1 for very high (sending) firing rates, although values of \alpha_i were chosen to be low so that EPSPs and IPSPs at individual synapses could summate to a certain extent across spikes (see also Wang, 1999). Sample trains of EPSPs and IPSPs are shown in Figure 3.2.
Figure 3.2: The $x$ and $s$ gating variables that mediate AMPA synaptic currents are shown in (A) and (B) for a random Poisson input spike train. The resultant EPSPs induced in a post-synaptic cell are shown in (C). (D)-(F) show the same information for a GABAergic input.
3.1.2 Firing-rate Adaptation and Synaptic Depression

The calcium-dependent potassium current \( I_{K(Ca)} \), also often referred to as the after-hyperpolarization current or \( I_{AHP} \), was responsible for fast firing-rate adaptation effects:

\[
I_{K(Ca)} = - g_{K(Ca)} \cdot s_{K(Ca)} \cdot (V - V_K)
\]

\[
dx_{K(Ca)}/dt = \sum_l \delta(t - t_l) - x_{K(Ca)}/\tau_{K(Ca)}
\]

\[
ds_{K(Ca)}/dt = \alpha_{sK(Ca)} \cdot x_{K(Ca)} \cdot (1 - s_{K(Ca)}) - s/\tau_{sK(Ca)}
\]

where \( g_{K(Ca)} \) is the maximal conductance (0.1 mS/cm²); \( s_{K(Ca)} \) is the dynamic adaptation gating variable, implemented similarly to the synaptic currents above as a system of two differential equations; \( V_K \) is the reversal potential for potassium (-90 mV); \( x_{K(Ca)} \) is a variable that increments instantaneously with each spike \( l \) and decays exponentially back to 0 with time constant \( \tau_{xK(Ca)} \) (0.2 ms), determining the rise time of \( s_{K(Ca)} \); \( \alpha_{sK(Ca)} \) determines the height of \( s_{K(Ca)} \) without dramatically affecting its time course (0.11/\( \tau_{sK(Ca)} \)); and \( \tau_{sK(Ca)} \) is the time constant that determines the decay rate of adaptation (80 ms). The parameters were fit to match the degree and time course of firing-rate adaptation measured in vivo for visual cortical neurons in anesthetized cats using intracellular recording techniques (Ahmed et al., 1998). As an excitatory neuron is driven to spike by its synaptic inputs, \( s_{K(Ca)} \) moves toward a value of 1. This gates open \( I_{K(Ca)} \), pushing \( V \) towards \( V_K \), and reducing the firing rate of the cell. If the synaptic drive is then removed, \( s_{K(Ca)} \) recovers exponentially back to 0 at the rate of \( \tau_{sK(Ca)} \), leaving a slow after-hyperpolarization of \( V \) until \( I_{K(Ca)} \) returns to 0. These points are depicted graphically in Figure 3.3.
Figure 3.3: (A) The impact of firing-rate adaptation on a cell's spiking responses when driven with a constant input current. Adaptation leads later spikes to be spread further apart relative to initial spikes, and the removal of current reveals a slowly recovering after-hyperpolarization. (B) shows the same cell and applied current as in (A), except with adaptation blocked. Spikes have been added to the voltage traces in order to aid visualization.
The implementation of synaptic depression was taken from Varela et al. (1999), for which parameters had been fit to intracellular recording data from V1 neurons (see also Varela et al., 1997). Recall from Section 3.1.1 above that synaptic depression is included as a dynamic scaling factor ($D$) on the outgoing AMPA and GABAa synaptic currents for each cell. At excitatory synapses (this included synapses from input neurons and those in the coupled pool), $D$ is actually composed of two individual scaling terms, one fast and one slow ($D_e = D_{e[fast]} \cdot D_{e[slow]}$). Inhibitory synapses, in contrast, have only one scaling term ($D_i$). The dynamics of each scaling term, like those of the synaptic currents and firing-rate adaptation, involve two differential equations:\(^{12}\):

$$\frac{dx_D}{dt} = \sum_j \delta(t - t_j) - x_D/\tau_{D}
$$

$$\frac{dD}{dt} = d^* \cdot x_D \cdot D_j + (1 - D_j)/\tau_D$$

where $x_D$ is a variable that is incremented instantaneously by 1 at the time of each pre-synaptic spike $l$ in neuron $j$ and recovers back to 0 with a time constant of $\tau_D$ (0.2 ms), determining the rate of decrease in $D_j$ immediately following each spike; $d^*$ is a parameter that determines how much $D_j$ decreases in response to each pre-synaptic spike $[d^*_{e[fast]} = \ln(0.78)/\tau_D, d^*_{e[slow]} = \ln(0.97)/\tau_D, d^*_{i} = \ln(0.94)/\tau_D$, where 0.78, 0.97, and 0.94 correspond to the values of the original $d$ parameters in the Varela et al. (1999) model]; and $\tau_D$ is the time constant determining how rapidly $D_j$ recovers back to 1 between pre-synaptic spikes ($\tau_{D_{e[fast]}} = 634$ ms, $\tau_{D_{e[slow]}} = 9300$ ms, $\tau_{D_i} = 1900$ ms).

---

\(^{12}\) The only alteration of the basic Varela et al. (1999) model was to apply decreases to the depression scaling terms gradually throughout the refractory period of each spike rather than instantaneously when the spike threshold was crossed. This protected against any artifactual influences of instantaneous changes in the scaling terms on the dynamics of neural synchrony; gradual changes applied during the time course of a spike are more likely to be the actual state of affairs in real neurons.
Figure 3.4: Dynamics of synaptic depression at excitatory versus inhibitory synapses. (A) shows that depression is stronger at excitatory ($D_e$) than at inhibitory synapses ($D_i$) for the same fixed firing rate. (B) shows the reductions in the heights of EPSPs due to synaptic depression at AMPA synapses, and (C) shows the same information for inhibitory synapses. Consistent with (A), reductions are more severe at excitatory synapses.
When a pre-synaptic neuron $j$ begins to spike at a high rate, $D_j$ scales down toward 0 with each spike. If spike activity ceases for a period of time, $D_j$ recovers exponentially back to 1. Figure 3.4 shows $D_j$ for excitatory versus inhibitory synapses over a train of pre-synaptic spikes and the resulting weakening of EPSPs and IPSPs.

### 3.1.3 Stimuli

Stimuli were administered to the network by generating random spikes independently in each of the 1000 input cells with Poisson characteristics for 500 milliseconds [mean input firing rate (±SD) = 30 ± 8 Hz]. Slightly different durations were used in some circumstances to promote consistency with particular experimental paradigms employed with humans and monkeys (reported below where relevant). Input firing rates were also sine-modulated in the 0-π range of the sine function over the stimulus duration (sample input spike trains are shown in Figure 3.5). This helped to capture the transient and non-uniform characteristics of visually evoked sensory inputs to visual cortical neurons.\(^\text{13}\) Stimuli were repeated up to 10 times in a stimulus block and were separated by a fixed inter-stimulus interval (ISI). The numbers of repetitions and the values of the ISIs were chosen to match particular experimental paradigms, and the values used in each simulation will be discussed below (ranging from 700 milliseconds up to 20 seconds). Each fixed ISI stimulus block was run ten different times to yield stable estimates of the critical dependent measures (firing rate and synchrony, discussed below) and to facilitate statistical comparisons across conditions.

\(^{13}\) Sine-modulated inputs were not critical for any of the effects that will be reported. They did, however, help to spread out the pattern of spikes over the time window of each stimulus, due to the interaction of input stimulation with synaptic depression and firing-rate adaptation.
Figure 3.5: A sample stimulus presented to the model. (A) shows the sine-modulated input spike trains across all 1000 excitatory input neurons (each dot corresponds to the occurrence of an individual spike), and (B) shows more detail for neurons 1-20.
3.1.4 Measuring Firing Rate and Synchrony

3.1.4.1 Firing Rate

In most simulations, firing rates were recorded for each cell in the network during each stimulus presentation by counting the number of spikes that occurred in the first 200 milliseconds of stimulation and dividing this count by the corresponding duration in seconds (0.2 s) to yield values in spikes/second (Hz). This interval was chosen to be comparable with that used to calculate firing rates in prior neural recording studies of repetition suppression (e.g. Li et al., 1993; Miller et al., 1993). In some cases, firing rates were calculated over the entire stimulus duration to promote consistency with particular studies (mentioned below where relevant).

3.1.4.2 Synchrony

A graded measure of neural synchronization, coherence \( \text{Coh} (\text{neuron}_1, \text{neuron}_2) \), was adapted from White et al. (1998) in order to evaluate the degree of synchrony between pairs of neural spike trains. White et al. (1998) introduced a simple method for calculating synchrony between two spike trains that normalized for differences in firing rate. Their basic approach, shown in Figure 3.6, involved: 1) replacing each spike in the two spike trains with square pulses of height 1.0 (from 0.0) and width equal to 0.2 of the period of firing \( T \) of the faster cell (assuming relatively stable firing rates); each square pulse was centered at the time of each spike, and 2) calculating the cross-correlation between the two trains of square pulses at zero time lag (i.e. the sum over the product of
the two pulse trains divided by the square root of the product of the sums of the individual trains):

\[
\text{Coherence (} n_1, n_2 \text{)} = \frac{\int_{t_0}^{t_f} P_{n_1} \cdot P_{n_2}}{\left( \int_{t_0}^{t_f} P_{n_1} \cdot \int_{t_0}^{t_f} P_{n_2} \right)^{1/2}}
\]

where \( n_1 \) and \( n_2 \) are the two neurons in the pair and \( P_{n_1} \) and \( P_{n_2} \) are the corresponding pulse trains (\( t_0 \) and \( t_f \) are the initial and final times). Coherence values for two spike trains range between 0 (no overlap of the square pulses) up 1 (complete overlap indicating perfect synchrony). Higher firing rates do not inflate this coherence measure because the widths of the square pulses shrink proportionately to the firing rates of the cells.
Figure 3.6: Simple method for calculating coherence between two spike trains used by White, Chow, Ritt, Soto-Trevino, and Kopell (1998). Two spike trains shown in (A) and (B) are replaced by square pulses with a width of 20% of the period of the fastest firing cell ($T_1$ here). The shaded area in D shows the overlap between the two square-pulse trains that is summed over time and used to calculate coherence. See text for more details. From "Synchronization and oscillatory dynamics in heterogeneous, mutually inhibited neurons," by J. White, C. Chow, J. Ritt, C. Soto-Trevino, and N. Kopell, 1998, *Journal of Computational Neuroscience*, 5, p. 7. Copyright 1998 by Kluwer Academic Publishers. Permission pending.
This basic coherence measure was adapted to handle time-varying firing rates by incorporating a simple dynamic estimate of firing rate for each cell that at any given time depended only on the spacing of the temporally adjacent spikes (either previous or subsequent). This modification was critical because without it, the systematic changes in firing rate across the network due to firing-rate adaptation and synaptic depression would artificially inflate the coherence values: If a single average firing rate is calculated for a cell over a duration when the rates are steadily attenuating, overall lower rates - and wider square pulses - will be assigned to the earlier spikes. Because most of the spikes in this case will be distributed at the beginning of the spike trains for a given stimulus, the coherence values will be spuriously inflated, particularly if the cells exhibit an initial burst of high firing rate at the onset of a stimulus (as they often do). The exact procedure for calculating the instantaneous firing rate of each cell during a single stimulus presentation was as follows: 1) a single value of firing rate was assigned to the midpoint of each inter-spike interval during the cell's spike train (simply 1 divided by the duration of the inter-spike interval), and 2) a single value of firing rate was then calculated at the time of each spike by linearly interpolating between the firing rates assigned to the preceding and subsequent inter-spike intervals. For cases in which a stimulus only elicited a single spike in a cell, its firing rate at all times was set equal to 1 divided by the stimulus duration. If only two spikes were elicited, the inter-spike interval between them was used to calculate a single firing rate at all times. The initial and final spikes in a train were assigned firing rates based on the first and last inter-spike interval, respectively. Coherence between two spike trains was then calculated by assigning square pulses to each spike with height of 1.0 (from 0.0) and width of 0.2 multiplied by the period of the
fastest firing cell at the time of the spike (the firing rate of the cell not owning a particular spike was interpolated from the rates previously calculated at its adjacent spikes). To insure that each spike in the train counted equally toward the coherence measure, the height of each square pulse in the individual pulse trains (initially 1.0) was then adjusted by dividing it by its own width (giving each pulse an area of 1.0). The height of each pulse in the product of the two trains was then set to 1.0/width of the thinner of the two individual overlapping pulses. The standard cross-correlation was then calculated using the pulse trains with the modified heights. Like the simpler coherence measure used by White et al. (1998), the modified coherence measure ranged between 0 (no overlap between the pulse trains) and 1 (complete overlap indicating perfect synchrony). Indeed, for periodically firing neurons, the two measures produce identical results. However, when firing rates vary widely during a stimulus such as when an entire population of neurons exhibits short bursts of high rate during which individual spike times are nevertheless asynchronous, the modified measure correctly gives low coherence values whereas the White et al. (1998) measure produces high values (see Figure 3.7).
Figure 3.7: A comparison between the simple measure of coherence used by White et al. (1998) and the modification used in the current simulations that involved a time-varying estimate of firing rate. Two sample spike trains are shown in (A) over a short time window. The pulse trains and overlap generated using time-varying firing rate are shown in (B) and the trains using fixed firing rate (as in White et al., 1998) are shown in (C). Short bursts of firing rate that are asynchronous give less overlap and lower coherence values using the method in (B) than in (C).
3.2 Simulation 1: The Effect of Stimulus Repetition on Firing Rate

This simulation evaluated the extent to which the basic model characterized in Section 3.1 above was capable of accounting for the magnitude and time course of short-term repetition suppression effects in the neural recording experiments of Miller, Desimone, and colleagues (e.g. Li et al., 1993; Miller et al., 1991, 1993) and the neuroimaging studies of Jiang et al. (2000) and Grill-Spector and Malach (2001). Given that the parameters of the short-term plasticity mechanisms in our model have been directly constrained by other sources of empirical data, success here will demonstrate that local adaptation and synaptic depression as previously observed are good candidate explanations of short-term repetition suppression effects in cortical cells.

3.2.1 Short-term repetition suppression in anesthetized monkeys (Miller, Gochin, & Gross, 1991)

Miller, Gochin, and Gross (1991) studied the effect of short-term stimulus repetition on neural responses in the inferotemporal cortex of macaque monkeys. Monkeys were either anesthetized (nitrous oxide) with their eyes focused on a target screen or were awake and trained to maintain fixation on the center of the target screen. Single common objects or photographs of common objects (e.g. stuffed animals, plastic toys, brushes, etc.) were presented repeatedly up to 9 times with a stimulus duration of 1 second, and repetitions were separated by a fixed interstimulus interval (ISI) of either 2, 4, 6, 12, or 20 seconds. Spike counts in single inferotemporal cortical cells were recorded during
each stimulus presentation. Changes in firing rate across repetitions for anesthetized monkeys are shown separately for each ISI in Figure 3.8A. Firing rates were re-centered in each condition around 0 spikes per second (Hz) by subtracting the initial firing rate from all values. The actual initial firing rates are given in parentheses, and the spontaneous or "baseline" firing rates are shown for reference with solid triangles. Firing rates decreased across repetitions, with greater decreases for shorter ISIs. In order to evaluate the model's ability to address these data, we similarly exposed it to a sequence of 9 stimulus repetitions with ISIs of 2, 4, 6, 12, and 20 seconds. As discussed above in Section 3.1.3, stimuli were 500 milliseconds in duration - similar to those used with monkeys, and firing rates were calculated over the first 200 milliseconds. Each ISI condition was repeated 10 times to yield stable estimates of firing rate. Figure 3.8B shows the changes in firing rate for the 250 excitatory cells in the model across each stimulus repetition and for each ISI, re-centered by subtracting the initial firing rate from each value (average initial firing rate of 31.6 Hz). Only the activities of the excitatory cells were included, as most (if not all) of the cells contributing to results in extracellular neural recording experiments are large pyramidal cells that are more easily isolable. However, the firing rates of the 50 inhibitory cells show a comparable pattern of change.
Figure 3.8: Effect stimulus repetition and inter-stimulus interval (ISI) on firing rates in monkey inferotemporal cortex (A) reported by Miller, Gochin, and Gross (1991) and in the model (B). The changes in firing rate that build up across repetitions in both cases are smaller with longer ISIs. From "Habituation-like decrease in the responses of neurons in inferior temporal cortex of the macaque," by E. Miller, P. Gochin, and C. Gross, 1991, *Visual Neuroscience*, 7, p. 360. Copyright 1991 by Cambridge University Press. Permission pending.
As expected, firing rates decreased across stimulus repetitions and decreases were greater for shorter ISIs. Indeed, the time course and magnitude of the firing rate changes across repetitions were quite comparable to the results of Miller et al. (1991) for ISIs of 2, 4, and 6 seconds (asymptoting at decreases of -12, -9, and -6, respectively). The results for ISIs of 12 and 20 seconds matched less well, and the pattern apparent at an ISI of 20 seconds in the Miller et al. (1991) data - a decrease followed by a later increase - may reflect some longer-term learning processes not incorporated in the model. Nevertheless, the overall interaction of repetition suppression with ISI is reasonably captured by the model. The statistical significance of the effects of stimulus repetition and ISI on firing rate was assessed in a repeated-measures ANOVA with repetition as a within factor and ISI as a between factor. There were highly significant main effects of repetition \( F(9, 405) = 371.04, p < .0005 \) and ISI \( F(4, 45) = 2751.71, p < .0005 \), as well as a highly significant interaction of repetition and ISI \( F(36, 405) = 47.20, p < .0005 \), indicating that the firing rate decreases that resulted from stimulus repetition were significantly larger for shorter ISIs. Individual repeated measures ANOVAs calculated for each ISI also showed that the effect of repetition was highly significant for each ISI \( p < .0005 \) for ISI's 2, 4, 6, and 12 seconds; \( p < .001 \) for an ISI of 20 seconds.
3.2.2 Short-term repetition suppression in awake monkeys performing a Delay Match-to-Sample (DMS) task (Miller, Li, & Desimone, 1993)

Several of the more detailed neural recording studies of short- and long-term repetition suppression effects in awake behaving monkeys have employed a Delay Match-to-Sample task (e.g. Li et al., 1993; Miller et al., 1993; Rainer & Miller, 2000). In this task, the monkey is exposed at the beginning of each trial to a sample stimulus and then to a series of individual stimuli, one of which is identical to the sample (the match). The monkey's job is to indicate the match by releasing a bar, ignoring any nonmatching stimuli that intervene (shown graphically in Figure 3.9).

Figure 3.9: Schematic diagram of the delay match-to-sample (DMS) paradigm used with macaque monkeys by Miller, Li, and Desimone (1993). A sample stimulus is presented for 500 ms, followed by between 0-4 distractor or "nonmatching" stimuli (each also presented for 500 ms and separated by 700-ms delays), and finally by the matching stimulus that required the monkey to respond by releasing a bar. From "Activity of neurons in anterior inferior temporal cortex during a short-term memory task," by E. Miller, L. Li, and R. Desimone, 1993, *Journal of Neuroscience, 13*, p. 1461. Copyright 1993 by the Society for Neuroscience. Permission pending.
In their study of short-term repetition suppression effects, Miller et al. (1993) presented each stimulus (sample, nonmatch, or match) for 500 milliseconds with ISIs of 700 milliseconds. Spikes in individual inferotemporal cortex neurons were recorded during each stimulus presentation, and firing rates were calculated over the first 200 milliseconds of the visually evoked responses (beginning at 75 milliseconds post-stimulus onset). Firing rates to stimuli eliciting significant match-nonmatch effects in the recorded cells are shown in Figure 3.10 A for different numbers of intervening nonmatch stimuli (0-4) between the sample and match stimuli. Firing rates to the match stimuli showed greater reductions relative to the sample with fewer intervening nonmatch stimuli. In addition, firing rates decreased slightly across several different intervening nonmatch stimuli.

The ability of the model to address these data was assessed by presenting two repetitions of a single stimulus (sample and match) with ISIs of 0.7, 1.9, 3.1, 4.3, and 5.5 seconds to simulate the amount of time that would elapse during different numbers of intervening nonmatch stimuli (0-4). The stimulus duration was 500 milliseconds (as in the experiment and previous simulation), and the firing rate was calculated over the first 200 milliseconds of the stimulus-evoked responses (i.e. starting at stimulus onset). While the model did not represent multiple stimuli, one can estimate responses to nonmatch values by assuming minimal overlap in the neural representations of different stimuli; under these circumstances, stimuli presented as nonmatch would be expected to yield firing rates comparable to the firing rates evoked by the sample stimulus. With this assumption, the firing rates to the sample stimulus have been substituted as rough estimates for nonmatch stimuli for ease of comparison with the match stimulus after
different numbers of intervening stimuli (i.e. after different ISIs). The average firing rates of the 250 excitatory cells to sample, "nonmatch", and match stimuli are presented in Figure 3.10 B. As with the empirical data, the firing rates of the model show the greatest decreases to match stimuli presented after fewer intervening nonmatch stimuli. In addition, the actual firing rate magnitudes of the model for sample and match stimuli are quite comparable to the recorded rates following each delay condition. The one obvious shortcoming of the model is its inability to address firing rate decreases to nonmatch stimuli. This might be improved in the future by representing several stimuli and allowing some partial overlap in the population of neurons that respond to each stimulus. Some partial firing rate decreases could then be observed across different stimuli due to the build up of adaptation and synaptic depression in neurons common to some or all of the stimulus representations (see Sohal & Hasselmo, 2000, for a similar account of decreased firing rates to nonmatch stimuli). Despite this shortcoming, the differences between match and nonmatch firing rates in the model are still quite comparable to the recorded rates after 0, 1, and 2 intervening nonmatch stimuli, only showing poor correspondence after 3 intervening stimuli (see Figure 3.11).
Figure 3.10: (A) Effect of repeating a stimulus (Sample to Match) on the firing rates of monkey inferotemporal cortex cells as a function of the number of intervening nonmatching stimuli in the DMS task, as reported by Miller, Li, and Desimone (1993). Responses to nonmatching stimuli are also shown. (B) Analogous results for the model. The difference between match and nonmatching stimuli decreased with more intervening stimuli. From "Activity of neurons in anterior inferior temporal cortex during a short-term memory task," by E. Miller, L. Li, and R. Desimone, 1993, *Journal of Neuroscience, 13*, p. 1465. Copyright 1993 by the Society for Neuroscience. Permission pending.
Figure 3.11: This figure re-graphs the differences between firing rates to match and nonmatch stimuli shown in Figure 3.10 for (A) the empirical data (Miller, Li, and Desimone, 1993) and (B) the model. These differences decrease similarly for both as a function of the number of intervening nonmatch stimuli. (A) also plots a dotted line corresponding to the monkeys’ level of correct performance in the DMS task, showing that performance is correlated with the difference between match and nonmatch firing rates. From "Activity of neurons in anterior inferior temporal cortex during a short-term memory task," by E. Miller, L. Li, and R. Desimone, 1993, *Journal of Neuroscience, 13*, p. 1467. Copyright 1993 by the Society for Neuroscience. Permission pending.

Repeated measures ANOVAs revealed highly significant main effects of repetition $[F(1, 45) = 1392.26, p < .0005]$ and number of intervening stimuli (ISI) $[F(4, 45) = 9.61, p < .0005]$, as well as a highly significant interaction indicating greater effects of repetition after fewer intervening stimuli $[F(4, 45) = ., p < .0005]$.

In order to further quantify the match suppression phenomenon, Miller et al. (1993) characterized the probability that a match stimulus would elicit weaker versus stronger firing rates relative to either sample or nonmatch stimuli. They developed an index to reveal the direction and magnitude of firing rate differences (sample vs. match or match vs. nonmatch) for each stimulus relative to the overall magnitudes of the firing rates.
rates: \( \frac{\text{rate}_{\text{Match}} - \text{rate}_{\text{Sample/Nonmatch}}}{\text{rate}_{\text{Match}} + \text{rate}_{\text{Sample/Nonmatch}}} \). A positive value would indicate a higher firing rate (averaged across recorded cells) to the match than to the sample/nonmatch for a particular stimulus, whereas a negative value would indicate a lower firing rate to the match. The distributions of these indices over stimuli that evoked significant match-nonmatch differences are shown in Figure 3.12 A and B, and the distributions over all stimuli that evoked above-baseline neural responses are shown in Figure 3.12 C and D. Figure 3.12 A and C show index values for match relative to nonmatch stimuli, and B and D show index values for match relative to sample stimuli. In all cases, most stimuli produced negative index values, a trend that was particularly marked for the set of stimuli yielding significant match-nonmatch differences and the plots comparing match and sample stimuli (Figure 3.12 A, B, and D). Figure 3.12 E plots the same index calculated for the model by taking each cell’s firing rates to sample and match in each run of each ISI condition. The probability distribution of the index values was then estimated by tabulating them in bins of 0.05 and dividing each bin count by the total number of values. While this distribution has been calculated over cells rather than stimuli, one might expect results to be similar for reasons that will be made clear shortly.
Figure 3.12: Distributions of repetition suppression indices calculated for the empirical data (A-D), as reported by Miller, Li, and Desimone (1993), and for the model (E). Negative values of the suppression index indicate greater firing rates to the nonmatch and sample stimuli than to the matching stimuli. (A) match-nonmatch effects for stimuli that evoked match-nonmatch differences, (B) match-sample effects for stimuli that evoked match-nonmatch differences, (C) match-nonmatch effects for stimuli that evoked above-baseline responses, (D) match-sample effects for stimuli that evoked above-baseline responses, (E) match-sample effects in the model. From "Activity of neurons in anterior inferior temporal cortex during a short-term memory task," by E. Miller, L. Li, and R. Desimone, 1993, *Journal of Neuroscience, 13*, p. 1464. Copyright 1993 by the Society for Neuroscience. Permission pending.
Note that the indices for the model are distributed similarly to the distributions calculated for the empirical data. Most index values are negative with a median of -0.11 and a standard deviation of 0.1. However, there are some positive values, as well (constituting approximately 5% of responses). Positive values are possible in the model for at least a couple of different reasons. Noise in the spike times of the input neurons could by chance lead some cells to get greater input on the match stimulus relative to the sample. It is also the case that randomness in the connectivity of the network leads some cells to receive a small number of excitatory inputs relative to the number of inhibitory inputs (the level of heterogeneity for these simulations was +/- 20%). When firing rate adaptation and synaptic depression lead to large reductions in the firing rates of all cells - including inhibitory cells, the ratio of excitation to inhibition can actually increase slightly for cells receiving greater inhibition, leading to small or modest increases in firing rate. This is reflected in Figure 3.13, where firing rate to the sample stimulus is plotted against the change in firing rate from sample to match (firing rate to the match, Stimulus 2, subtracted from the firing rate to the sample, Stimulus 1). Data points occur at fixed intervals because firing rates all occur at multiples of 5 Hz when calculated over a 200 millisecond time window. Note that firing rate increases tend to occur for lower initial firing rates, consistent with an initial low ratio of excitation to inhibition that increases slightly with repetition due to reduced inhibition. Noisy inputs may contribute to a certain extent, as well. Firing rate decreases, on the other hand, tend to occur for higher initial firing rates. Indeed, the change in firing rate is well correlated with the initial firing rate \[ R(12498) = .6, p<.0005 \].
Figure 3.13: The degree of repetition suppression in the model from Stimulus 1 to Stimulus 2 was larger for larger Stimulus 1 firing rates. The positive suppression indices observed in Figure 3.12 were for smaller firing rates to Stimulus 1. This indicates that for the model, firing rate decreases will be largest for cells that respond most vigorously to a stimulus, leading to decreased stimulus selectivity.

This also helps to explain the gap between positive and negative index values in Figure 3.12 E, because few firing rate increases are occurring at higher firing rates, leading to fewer small positive index values. Large decreases in firing rate, along with some small to moderate increases appear to be consistent with short-term plasticity mechanisms such as firing-rate adaptation and synaptic depression.

These analyses raise the question: Do stimuli in the Miller et al. (1993) study that give rise to positive index values elicit lower average firing rates than those that elicit negative values? The answer appears to be 'yes'. Figure 3.14 A shows the firing rate responses to stimuli that give rise to greater nonmatch responses than match responses, and Figure 3.14 B shows firing rates to stimuli that induce greater match responses.
Figure 3.14: Match suppression effects, as reported by Miller, Li, and Desimone (1993), were observed for stimuli that evoked larger firing rates (A), whereas match enhancement effects were observed for stimuli that evoked smaller firing rates (B). This is potentially analogous to the results shown for the model in Figure 3.13. From "Activity of neurons in anterior inferior temporal cortex during a short-term memory task," by E. Miller, L. Li, and R. Desimone, 1993, *Journal of Neuroscience, 13*, p. 1466. Copyright 1993 by the Society for Neuroscience. Permission pending.

Note that firing rates are larger overall when nonmatch and sample responses exceed match responses. For comparison, the model produces larger match than sample responses (match enhancement) when the average firing rate to the sample is 22.48 Hz (SE = 0.25), whereas it produces larger sample than match responses (match suppression) when the average firing rate to the sample is 33.61 Hz (SE = 0.09). These values are reasonably close to the firing rates evoked by sample stimuli in Figure 3.14 A and B. It appears plausible then that the reason that some stimuli yield positive index values is that they induce low initial firing rates in the recorded cells; the neurons that happened to be
recorded for such stimuli were not particularly well-tuned for them, whereas they might in principle be quite responsive if the right stimulus were chosen. This is likely for the Miller et al. (1993) study because each cell that was recorded was only exposed to 6 different stimuli (chosen randomly from a set of 500 stimuli), and each stimulus was only presented during the recording of 1 to 2 different cells (although it was presented many times to those cells). The model may produce a similar distribution of indices when calculated across cells (as opposed to stimuli) for the same basic reasons: Some cells are responsive to the presented stimulus and others are less responsive. The model predicts that if single stimuli were presented during the recording of many different cells, the distribution of indices when calculated across cells (within each stimulus) should look much like Figure 3.12 E. It further predicts that the change in firing rate between sample and match should be correlated with the firing rate to the sample (Figure 3.13). While these predictions have not yet been examined very directly, Miller et al. (1993) did evaluate the relationship between stimulus tuning and the difference between match and nonmatch responses in the subset of cells that were stimulus-selective and sensitive to match/nonmatch status (N=64 out of 141 total). For 67% (43/64) of these cells, match-nonmatch differences were similar across their corresponding 6 stimuli, rank ordered from best to worst using the firing rate to each presented as the sample stimulus (an example of one such cell is shown in Figure 3.15 A).
Figure 3.15: Stimulus-selectivity in individual IT cells as a function of short-term stimulus repetition (Match vs. Nonmatch), as reported by Miller, Li, and Desimone (1993). Repetition suppression effects were not largest for poor stimuli (as defined by the firing rate elicited to each stimulus as Sample). Cells either showed comparable decreases for all stimuli or showed disproportionately large decreases for the best stimuli. This indicates that stimulus repetition led to overall decreases in stimulus selectivity. From "Activity of neurons in anterior inferior temporal cortex during a short-term memory task," by E. Miller, L. Li, and R. Desimone, 1993, Journal of Neuroscience, 13, p. 1463. Copyright 1993 by the Society for Neuroscience. Permission pending.

For most of the remaining cells (15/21), the match-nonmatch differences were disproportionately large to the best stimulus, other cells (see Figure 3.15 B for an example). This means that, if anything, repetition leads to a decrease in stimulus selectivity (based on firing rate) rather than an increase. For the opposite to be true, the greatest decreases would need to occur for the worst stimuli. While more work remains to be done on this issue, the current results are problematic for accounts of short-term
behavioral priming effects that rely on a short-term neural "sharpening" of stimulus selectivity despite overall reductions in firing rates. However, this does not necessarily rule out the possibility of such a mechanism for behavioral priming effects observed at longer time scales. These issues are discussed in more detail in Appendix A.

3.2.3 Short-term repetition suppression in human adults performing Delay Match-to-Sample (Jiang, Haxby, Martin, Ungerleider, & Parasuraman, 2000)

Jiang et al. (2000) used functional magnetic resonance imaging (fMRI) to study the changes neural activity (actually hemodynamic responses) that underlie human performance of the DMS task. Experimental participants were shown a sample face stimulus at the beginning of a trial and then were asked to press a button each time they saw a face that matched the sample. Target match faces and distractor faces were presented anywhere from 1 up to 5 times within a trial in an unpredictable sequence. They found that repetition of both targets and distractors led to reduced hemodynamic responses in posterior cortical regions (e.g. ventral temporal and parietal cortices), whereas repetition had little effect on activity in more frontal regions (frontal/insular cortex, as well primary and supplementary motor regions). In addition, both posterior and frontal regions showed enhanced responses to target faces. Figure 3.16 A and B show hemodynamic responses to within-trial repetitions 1 through 5 in ventral temporal versus frontal/insular cortex. Hemodynamic responses have been expressed in terms of proportion change relative to non-repeated distractor responses (e.g. -0.2 = 20% decrease from the average non-repeated distractor response).
Figure 3.16: Repetition suppression effects observed in an fMRI study of a delay match-to-sample (DMS) task in humans, as reported by Jiang, Haxby, Martin, Ungerleider, and Parasuraman (2000) (A, B), and similar effects in the model (C). Short-term repetition of targets and distractors led to decreased hemodynamic responses in ventral temporal cortices (A) but not in frontal/insular cortices (B). Decreased activity in the model was comparable in magnitude and time course to that observed for distractor stimuli in (A). From "Complementary neural mechanisms for tracking items in human working memory," by Y. Jiang, J. Haxby, A. Martin, L. Ungerleider, and R. Parasuraman, 2000, Science, 287, p. 644. Copyright 2000 by the American Association for the Advancement of Science. Permission pending.

For comparison, Figure 3.16 C shows the change in the model's spiking responses relative to its initial firing rate when exposed to 4 repetitions with the same stimulus duration (2 seconds) and inter-repetition interval (4 seconds) used by Jiang et al. (2000). This comparison is reasonable if one assumes that the hemodynamic response is directly proportional to spiking responses and that the neural representations of different stimuli
share very few neurons\textsuperscript{14}. Interestingly, the model's responses quite accurately reproduce the time course and magnitude of responses in ventral temporal cortex to repeated distractor stimuli (they appear to fall within the standard error of the mean for each data point). A repeated measures ANOVA showed that the decreases across repetitions were highly significant [F(3, 27) = 703.48, p < .0005]. However, the current model has no ability to address the enhanced responses to target faces or the sustained activity in frontal/insular areas. Following Jiang et al. (2000) and others (e.g. Miller, Erickson, & Desimone, 1996), we would suggest that frontal/insular regions actively maintain representations of the target face, providing a non-decrementing top-down input that biases and enhances perceptual processing of the target face, eliciting elevated activity in posterior regions when a match occurs. Future simulations are planned that incorporate frontal-like components capable of sustaining activity across delays in which direct input is absent (e.g. Gutkin, Laing, Colby, Chow, & Ermentrout, 2001; Wang, 1999). For the time being, however, the focus is on understanding the repetition-related decreases in posterior cortical regions. It is also important to note the ways in which these results differ from those of Miller et al. (1993) (see Section 3.2.2). As discussed in some detail by Miller and Desimone (1994; Miller, Erickson, & Desimone, 1996), repeating distractor stimuli between the sample and match stimuli can dramatically alter repetition-related changes in neural responses. In particular, when distractors are repeated, the dominance of match suppression gives way to a more even mixture of firing rate increases and decreases. Sustained delay-period activity in prefrontal cortex is also more

\textsuperscript{14} While some recent findings suggest that local field potentials (LFPs) - and hence population synaptic currents (Mitzdorf & Singer, 1979; Mitzdorf, 1987) - are more tightly correlated with the hemodynamic response in fMRI than are spiking responses (Logothetis, Pauls, Augath, Trinath, & Oeltermann, 2001), examination of summed synaptic currents in the model yields comparable results, perhaps because the current model - unlike a real cortical region - has few background inputs unrelated to stimulus processing.
readily observed under these circumstances. The earlier studies that employed DMS tasks (e.g. Li et al., 1993; Miller, Li, & Desimone, 1991, 1993) did not repeat distractors between sample and match stimuli, permitting the monkeys to perform the task on the basis of simple repetition (Miller & Desimone, 1994). Our account addresses only the automatic and short-term repetition-related decreases that are present across these circumstances in more posterior cortical regions. Under task conditions that do not emphasize working memory requirements (i.e. tasks not requiring the explicit detection of contingencies across trials), it may be possible to observe similar automatic and short-term decreases in more frontal, motor regions (e.g. Karni, Meyer, Jezzard, Adams, Turner, & Ungerleider, 1995). Mechanisms by which firing-rate adaptation and synaptic depression could be reduced to allow non-decrementing responses in certain cortical regions and task conditions will be discussed further in Simulations 6-8.

3.2.4 Short-term repetition suppression in human adults during an fMR-Adaptation paradigm (Grill-Spector & Malach, 2001)

Some recent fMRI studies have utilized a technique referred to as "fMR-Adaptation" in order to examine the putative overlap of different types of neural representations in simple identification, discrimination and working memory tasks (e.g. Grill-Spector, Kushner, Edelman, Avidan-Carmel, Itzchak, & Malach, 1999; Grill-Spector & Malach, 2001; Naccache & Dehaene, 2001). This method generally involves an "adaptation" phase during which stimuli are repeatedly presented within a short period of time, and a testing phase during which a stimulus property of interest (e.g. size, position, orientation, etc.) is altered, and the degree of recovery of neural activity is assessed for task-related cortical regions. It has been argued that the degree of recovery observed in a cortical
region is directly related to that region's importance in encoding the altered stimulus property, potentially allowing researchers to use this method to probe coding properties of different cortical regions throughout the brain (Grill-Spector & Malach, 2001). For the current context, we will focus on one particular study in which only the adaptation phase was performed.

Grill-Spector & Malach (2001) exposed adult participants to a series of adaptation trials, each involving 32 serially presented visual pictures of common animals, objects, and faces (stimulus duration of 875 milliseconds, ISI of 125 milliseconds). In some trials, the same picture was presented for all 32 stimuli, whereas for other trials, larger numbers of pictures were used to create several different repetition conditions (1, 2, 4, 8, and 32 stimuli). When multiple stimuli were used, they were cycled through repeatedly in the same order, yielding fixed inter-repetition intervals (e.g. 4 stimuli: "41234123..."). Because different numbers of stimuli were repeated in this fashion over the same absolute period of time (32 seconds), conditions with more stimuli involved fewer repetitions separated by longer intervals (see Figure 3.17 for a graphical depiction of the different repetition conditions). Subjects were requested to name each object covertly, and hemodynamic responses were recorded for the different repetition conditions; these conditions were alternated with 16-second periods of mean luminance blanks, randomly oriented triangles or highly scrambled pictures.
Figure 3.17: Schematic diagram of the fMR-Adaptation paradigm used by Grill-Spector and Malach (2001). Scans were for a total of 32 seconds during which 32 picture stimuli were presented at 1 per second. Different numbers of picture stimuli were presented repeatedly during the scan (1, 2, 4, 8, and 32) to yield different amounts of fMR-Adaptation. One stimulus presented repeatedly over the 32 seconds corresponded to 32 identical repetitions, whereas 32 stimuli presented over the 32 seconds implied a single presentation for each stimulus. From "fMR-adaptation: a tool for studying the functional properties of human cortical neurons," by K. Grill-Spector and R. Malach, 2001, Acta Psychologica, 107, p. 298. Copyright 2001 by Elsevier Science B.V. Permission pending.

Results for cortical regions in which visual objects yielded larger hemodynamic responses than the textured and scrambled images are shown in Figure 3.18 A. These regions included several ventral visual areas (the lateral aspect of the occipital lobe, the posterior inferotemporal sulcus, and the posterior to mid-fusiform gyrus; also referred to as the LOC or lateral occipital complex), a dorsal visual focus in the intraparietal sulcus, and a small region within the collateral sulcus. The dark-shaded regions in Figure 3.18 A show the different repetition conditions (labels corresponding to the number of stimuli
used). Perhaps not surprisingly, the smallest hemodynamic response was observed in the 1-stimulus condition (32 repetitions of the same stimulus) and the largest response was observed in the 32-stimulus condition (each stimulus presented once), with the remaining responses ranging in-between by rank order of the number of stimuli used (1 < 2 < 4 < 8 < 32).

This same stimulus paradigm was approximated in the model by assuming that the neural representations of the 32 stimuli overlapped minimally (see also Sections 3.2.2 and 3.2.3). Stimulus durations were 875 milliseconds, and ISIs were 125, 1125, 3125, 7125, and 31125 milliseconds to approximate the 1-, 2-, 4-, 8-, and 32-stimulus conditions, respectively. Firing rates were estimated over even intervals of 4 seconds within each 32-second period, yielding 8 firing rate values in each of the five repetition conditions. For conditions involving multiple stimuli, the firing rate for the single "stimulus" represented by the network was substituted for intervening stimuli in the average firing rate calculations. Results in terms of average firing rate relative to that with no stimulus (i.e. 0 Hz) are plotted for each repetition condition in Figure 3.18 B; conditions have been arranged similarly to the experimental results for ease of comparison (although they were, in fact, run completely independently). As with the simulation in Section 3.2.3, this comparison assumes that the hemodynamic response is directly proportional to neural activity (either firing rates or synaptic potentials).
Figure 3.18: fMR-Adaptation effects as a function of time for the different repetition conditions used in Grill-Spector and Malach (2001). (A) The greatest hemodynamic response decreases in humans were observed in the 1-stimulus condition, and the weakest decreases were observed in the 32-stimulus condition. (B) The model produced similar decreases in average firing rate. See text for explanation. From "fMR-adaptation: a tool for studying the functional properties of human cortical neurons," by K. Grill-Spector and R. Malach, 2001, *Acta Psychologica*, 107, p. 303. Copyright 2001 by Elsevier Science B.V. Permission pending.

The rank order of the different repetition conditions and the relative magnitudes of the rates show good correspondence to the hemodynamic responses observed in Grill-Spector and Malach (2001). Notice also that, similarly to the empirical data, dynamic within-
condition changes are observed in the 1-, 2-, and 4-stimulus conditions relative to the 8- and 32-stimulus conditions, although they are somewhat faster and sharper than the hemodynamic responses. A more detailed comparison of the relative magnitudes of activity in different conditions is shown in Figure 3.19, averaged over the 32-second durations. Results are expressed in terms of the number of stimulus repetitions per 32 seconds (or per 16 seconds) and the ratio of the average activity under repetition to the magnitude of activity when stimuli are presented only once (also referred to as the "adaptation ratio"). The number of stimulus repetitions here refers to the number of identical stimulus presentations past the first (i.e. 32 presentations of the same stimulus = 31 repetitions). Figure 3.19 B shows that the model's responses are quite comparable in time course and magnitude to the average hemodynamic responses shown in Figure 3.19 A. In the 1-stimulus (31 repetition) condition, responses are approximately 40% of those in the 32-stimulus (0 repetition) condition. Both the model's responses and the empirical data are also slightly bowed below a linear trend, obeying what appears to be a negative exponential function with increasing number of repetitions. Repeated-measures ANOVAs showed that the model's firing rates decreased significantly as a function of repetition \[F(4, 45) = 2668.41, p < .0005\].
Figure 3.19: Average activity decreases in the fMR-Adaptation paradigm for the different repetition conditions. The adaptation ratio corresponds to the hemodynamic response (or firing rate) when normalized by the responses in the 32-stimulus condition. Results are plotted as a function of the number of repetitions used in each condition rather than as a function of number of stimuli; the data have been recalculated as the number of repetitions per 16 seconds in order to show the exponential shape of function more clearly. (A) shows results for the human data in Grill-Spector and Malach (2001), and (B) shows the same results calculated for the model. From "fMR-adaptation: a tool for studying the functional properties of human cortical neurons," by K. Grill-Spector and R. Malach, 2001, *Acta Psychologica, 107*, p. 305. Copyright 2001 by Elsevier Science B.V. Permission pending.
3.2.5 Comment

The simulations presented in Sections 3.2.1 though 3.2.4 demonstrate that a relatively simple neural model with implementations of firing-rate adaptation and synaptic depression that are directly constrained by independent sources of neurophysiological data are capable of accounting for the time course, magnitude, and various other properties of short-term repetition suppression observed in extracellular neural recording experiments with monkeys and functional neuroimaging experiments with adult human participants. Several of the simulations required an assumption of minimal overlap between the neural representations of different stimuli. This assumption can be relaxed in future simulations by incorporating differentially interconnected groups of neurons that represent different stimuli, allowing a more veridical implementation of the experimental paradigms utilized in the various empirical studies (see also Sohal & Hasselmo, 2000).

Given that both firing-rate adaptation and synaptic depression can be suppressed by neuromodulators such as acetylcholine and norepinephrine, the current account makes the strong prediction that cholinergic and noradrenergic agonists should reduce short-term repetition suppression in neural recording and neuroimaging experiments (see also Simulation 8). A previous neural recording study by Miller and Desimone (1993) showed that the muscarinic cholinergic receptor antagonist scopolamine does not reduce short-term repetition suppression effects; this finding is consistent with the current account because cholinergic agonists (e.g. muscarine, carbachol) rather than antagonists are predicted to reduce repetition suppression. The model also makes predictions about the effects of NMDA-receptor blockers (e.g. APV) on short-term versus long-term repetition suppression and behavioral priming: Short-term effects on the order of seconds
should be spared, whereas longer-term effects on the order of minutes, hours, or days should be diminished. This follows from the assumption that longer-term plasticity mechanisms such as LTP and LTD are responsible for the longer-lasting neural and behavioral changes, and these mechanisms are disrupted by antagonists of post-synaptic NMDA receptors (e.g. Artola & Singer, 1987; Kanter & Haberly, 1990; Morris, Anderson, Lynch, & Baudry, 1986). Since short-term plasticity mechanisms of firing-rate adaptation and synaptic depression do not depend on NMDA receptors, short-term repetition suppression and behavioral priming effects should be relatively normal.

An account of short-term, as well as long-term repetition suppression effects has previously been proposed by Sohal and Hasselmo (2000). Their account of short-term repetition suppression effects is quite similar to the current proposal in that they rely on firing-rate adaptation due to calcium-activated potassium currents, and as such, it makes many of the same predictions. However, the time constant of decay on the calcium current that they included is not slow enough to address many of the effects that we have considered here, a point that they acknowledge ($\tau = 1$ second: Schwindt, Spain, Foehring, Stafstrom, Chubb, & Crill, 1988), nor were the relevant parameters controlling the magnitude of adaptation constrained directly by independent sources of data. They also used firing-rate style neurons in their simulations. Results from Appendix A would therefore suggest that their model predicts poorer rather than better identification performance with short-term repetition suppression. Finally, they did not incorporate synaptic depression, which of the short-term plasticity processes in the current model is clearly the more important of the two for addressing the slow buildup of firing rate decreases across repetitions (due to the slow recovery time constants of 0.6 and 9.3
seconds; see Section 3.1.2). Even if we had included the slower potassium current that depends on sodium rather than calcium (average time constant of 10 seconds, Sanchez-Vives et al., 2000), synaptic depression tends to have a stronger effect on firing rates due to its multiplicative rather than additive nature\(^\text{15}\). Nevertheless, future experiments might be able to distinguish which of the two processes is more important for actual neural activity decreases by means of intracellular recording techniques in vivo: Firing-rate adaptation due to the activation of potassium currents should elicit extended hyperpolarization of the membrane potential and changes in input resistance following periods of intense spiking whereas synaptic depression should not cause hyperpolarizations or input resistance changes (see Appendix D).

### 3.3 Simulation 2: The Effect of Stimulus Repetition on Synchronization

*Simulation 1* (Section 3.2) demonstrated that the short-term plasticity mechanisms of firing-rate adaptation and synaptic depression (primarily synaptic depression) account nicely for the time course and magnitude of short-term repetition suppression effects in extracellular recording and functional neuroimaging experiments. Following this, *Simulation 2* evaluates the basic hypothesis that these short-term plasticity mechanisms

---

\(^{15}\) When I have added an implementation of the Na\(^+\)-activated K\(^+\) current that is fit to the contrast adaptation data of Sanchez-Vives et al. (2000) recorded in V1 neurons, its impact on firing-rate decreases is substantially weaker than that of synaptic depression, and by itself, it does not appear to account well for the magnitudes of repetition suppression that are observed in neural recording and neuroimaging experiments.
simultaneously enhance spike synchronization as they reduce neural activity. As discussed in Section 2.2, theoretical studies of neural synchronization predict that lower firing rates tend to aid spike synchrony in networks of excitatory and inhibitory neurons. Since firing-rate adaptation and synaptic depression promote lower firing rates across repetitions, their role in altering spike synchrony was evaluated by detecting changes in a simple measure of synchrony (coherence) computed over all pair-wise combinations of cells (see Section 3.1.4.2). The robustness of changes in coherence to factors that may potentially disrupt spike synchrony such as input heterogeneity, synaptic delays, and interactions with other cortical regions is evaluated, and the decay of such changes across different inter-stimulus intervals (ISIs) is evaluated against the decay of short-term priming effects studied by McKone (1998) over similar time intervals.

3.3.1 The effect of stimulus repetition on spike synchronization and its tolerance to input heterogeneity

In this simulation, the stimulus duration and method of calculating firing rates were the same as for simulations in Sections 3.2.1 and 3.2.2 that were used to address repetition suppression in neural recording experiments (stimulus duration=500 ms; firing rates calculated over first 200 ms). There were 10 stimulus repetitions in each stimulus block, and the ISI between repetitions was held at 1 second. Input heterogeneity was manipulated for the 250 excitatory cells and 50 inhibitory cells as discussed in Section 3.1, with 4 different gradations of heterogeneity (+/- 0, 20, 40, and 60% of the average number of excitatory inputs). The average firing rates for the excitatory and inhibitory cells are shown at each of the 10 stimulus repetitions in Figure 3.20 A. Only the results for the +/- 0% heterogeneity condition are shown, as the average firing rates (though not
the variances) were comparable across different heterogeneity conditions. The standard deviation of firing rates calculated over the 250 excitatory cells for the 10 different runs of each heterogeneity condition is shown for each stimulus repetition in Figure 3.20 B\textsuperscript{16}. As stimuli are repeated, both the means and the standard deviations of the firing rates for the excitatory cells (as well as for the inhibitory cells) are reduced by the short-term plasticity mechanisms. This is predominantly the impact of synaptic depression, because the calcium-dependent potassium currents used here for firing-rate adaptation decay nearly completely by 250 milliseconds - much faster than the 1 second ISI. The compression of the variability in the firing rates is intuitively simple: Synaptic depression is implemented as a multiplicative scaling factor on the synaptic currents, and large synaptic currents are scaled down more by this factor than are smaller currents (50\% of 100 is 50 - a difference of 50, whereas 50\% of 1 is 0.5 - a difference of 0.5). Synaptic depression actually depends on pre-synaptic transmitter release rather than the post-synaptic synaptic currents, per se, and so it possible in principle for it to have a larger impact on smaller rather than larger synaptic currents (e.g. if the small synaptic currents are summed over a small number of inputs with high firing rates or high release probability, compared to summing over a large number of inputs with low rates or release probability). However, when many inputs of different rates are randomly connected to a cell, this possibility becomes less likely. Repeated measures ANOVAs showed that, as in previous simulations, average firing rates of the excitatory cells showed highly significant decreases across stimulus repetitions \([F(9, 324) = 1585.19, p < .0005]\).

\textsuperscript{16} As changes in synchrony in the excitatory cells are our primary concern here for the speed of signal propagation throughout a series of cortical regions (the outputs to other cortical regions are exclusively excitatory: White, 1989), the focus of the analyses will be mostly on these cells.
Figure 3.20: Both the means (A) and the standard deviations (B) of firing rates in the model decrease as a function of stimulus repetition. (B) also shows that increases in input heterogeneity from 0 to 60% of the mean number of excitatory inputs leads to increases in the standard deviation of firing rates.
The standard deviations in firing rate decreased across repetitions \([F(9, 324) = 986.30, p < .0005]\), and they were also larger overall for greater levels of input heterogeneity \([F(3, 36) = 4123.46, p < .0005]\). Additionally, there was a highly significant interaction of stimulus repetition and level of input heterogeneity \([F(27, 324) = 139.36, p < .0005]\), indicating that decreases in the standard deviations of firing rates were larger for greater levels of input heterogeneity.

Given these changes in firing-rate, how does synchrony change across stimulus repetition for different levels of input heterogeneity? Figure 3.21 shows the concomitant changes in average coherence among the excitatory cells across repetitions for each level of heterogeneity. As discussed in Section 3.1.4.2, a coherence value was calculated for each pair-wise combination of cells for each stimulus. Only spikes that occurred between 20 milliseconds post-stimulus onset and the offset of the stimulus were included in the calculations because the initial burst of activity would have artificially inflated the coherence values across repetitions (the onset of spiking was roughly the same for each stimulus, but as the rates decreased across repetitions, the precise onset of the first spike would contribute relatively more to the total coherence value). The values were then averaged within groups of pairings (e.g. excitatory to excitatory, inhibitory to inhibitory, excitatory to inhibitory) and across the 10 runs of each level of heterogeneity. Coherence values ranged from 0 (no overlap of the corresponding pulse trains) to 1 (identical pulse trains). Coherence values of cells with themselves were excluded from the averages, and each pair-wise combination was included only once. Coherence values for all pair-wise combinations of the 1000 input neurons are shown for reference, as each input neuron fired independently of all others with Poisson characteristics.
Figure 3.21: Coherence values among the 250 excitatory cells in the model as a function of stimulus repetition and level of input heterogeneity (0-60%). Repetition leads to increases in coherence for all of the heterogeneity conditions, but the increases are smaller for larger values of heterogeneity. Error bars indicate the standard error of the mean for each repetition/heterogeneity combination.

Coherence values increased across stimulus repetitions for all of the heterogeneity conditions, although the changes were considerably smaller for larger values of heterogeneity. Repeated measures ANOVAs showed highly significant main effects of stimulus repetition \( [F(9, 324) = 19.49, p < .0005] \) and heterogeneity \( [F(3, 36) = 248.03, p < .0005] \) on coherence, as well as a highly significant interaction of repetition and heterogeneity \( [F(27, 324) = 2.96, p < .0005] \), indicating larger increases in coherence across repetitions for weaker heterogeneity. It is worth mentioning, though, that increases were highly significant even for the largest value of heterogeneity \( [+/- 60\%: \text{initial firing rate} = 29.66 +/- 19.49 (SD)] \) \( (p<.0005 \text{ for } +/- 0, 20, \text{ and } 60\%; p<.01 \text{ for } +/- 40\%) \).
A more detailed view of the changes in coherence is possible by examining the plots of the spike times for each cell in the coupled pool. Figure 3.22 shows the equivalent of raster plots for the 250 excitatory and 50 inhibitory cells in the model at the 1st, 2nd, 5th, and 10th stimulus repetitions for the +/- 20% heterogeneity condition (firing rates to the 1st stimulus ranging between 10 and 60 Hz with a mean of ~ 30 Hz). Time along the x-axes is expressed in milliseconds during the stimulus presentation (0 to 500 ms + 1000 ms of ISI). The y-axes give the index number for each of the 300 cells (labeled 1001-1250 for the excitatory cells, 1251-1300 for the inhibitory cells; cells 1-1000, shown earlier in Figure 3.5, correspond to each of the input neurons). Each spike time for each cell during the stimulus duration is indicated by a small black dot. It is clear from Figure 3.22 A (Stimulus 1) that all of the cells fire less randomly during the presentation of the first stimulus than one might expect. The cells fire short high-rate bursts during the beginning of the stimulus and fire somewhat more randomly after that. This is partially due to the fact that all of the cells start each stimulus near similar values of voltage and are driven to spike strongly initially due to strong synaptic currents that have not yet been attenuated due to synaptic depression. However, it is also likely due to the dynamical interactions of the spiking activity with short-term plasticity. The high-rate bursting shown here during the first stimulus is similar in nature to that observed by Tsodyks, Uziel, and Markram (2000; Loebel & Tsodyks, 2002) in their study of networks of integrate-and-fire (IAF) neurons with fast synaptic depression. Spike times during the burst are asynchronous due to the high firing rates that result from runaway positive feedback interactions between the excitatory cells; activity is then abruptly terminated by the very rapid buildup of synaptic depression and firing-rate adaptation.
Figure 3.22: Raster plots on a sample run of the model for Stimulus 1 (A), Stimulus 2 (B), Stimulus 5 (C), and Stimulus 10 (D). The spike times of the 300 cells (1000-1250 excitatory and 1251-1300 inhibitory) over the course of each stimulus (1000-1500 ms) are indicated in the plots by individual black dots. It is apparent that large changes in spike synchronization occur from Stimulus 1 to Stimulus 2 (A vs. B). High levels of synchronization are maintained across stimulus repetitions (B-D).

Figure 3.22 B-D, in contrast, show a high degree of synchrony for individual spikes throughout the duration of the stimulus presentation (2nd, 5th, and 10th repetitions, respectively). It appears, however, that the benefit of this coordinated activity over and above that elicited by the first stimulus would not be observed until the later spikes in the train (during 100-500 ms post-stimulus onset). This issue will be taken up in greater detail in *Simulation 4 (Section 3.5).*
3.3.2 The impact of synaptic delays and interactions with other cortical regions on repetition-related changes in spike synchronization

Communication between real neural cells is not instantaneous. Voltage impulses must travel down the length of the axon from the soma (where action potentials are initiated), and then transmitter must diffuse across the synaptic cleft to bind with the post-synaptic neuron's AMPA, NMDA, or GABA receptors, giving rise to a synaptic potential. All of these processes take time. In order to examine the impact of including fixed synaptic delays on repetition-related changes in synchronization, we compared the coherence values calculated for the model when synaptic delays of 0, 1, and 2 milliseconds were used (comparable to those used in other models of neocortex, e.g. Chawla, Lumer, & Friston, 1999). The simulation was similar in all other respects to that discussed above in Section 3.3.1, except that heterogeneity was not varied (fixed at +/- 0%)\(^{17}\). Figure 3.23 A shows the changes in coherence values for the excitatory cells across stimulus repetitions for each synaptic delay condition, and Figure 3.23 B shows changes in coherence among the inhibitory cells. Firing rates for all three conditions were comparable to those shown above in Figure 3.20. Coherence values among the excitatory cells increased significantly across repetitions for each value of synaptic delay (\(p<.0005\) for delays of 0, 1 ms; \(p<.004\) for 2-ms delay). Coherence values were also larger for a delay of 0 ms compared to delays of 1 and 2 ms (\(p<.0005\) for 0 vs. 1 ms and 0 vs. 2 ms), with no significant difference between delays of 1 and 2 ms (\(p>.6\)). While the repetition-related changes in coherence appeared larger for a delay of 0 ms, the interaction between repetition and synaptic delay did not reach significance (\(p>.1\)).

\(^{17}\) Similar results have been obtained with +/- 10 and 20% heterogeneity and synaptic delays of 1 ms.
Figure 3.23: The impact of synaptic delays (0-2 ms) on repetition-related changes in coherence. (A) shows changes in coherence among the excitatory cells, whereas (B) shows the changes among the inhibitory cells. (A) shows that coherence increases are robust for each synaptic delay condition, whereas (B) shows a more complicated relationship for the inhibitory cells. See text for explanation.

When one considers the coherence values among the excitatory versus the inhibitory cells (compare Figure 3.23A and B), it is clear that coherence is lower overall for the inhibitory cells. This finding is also borne out in the plots of spike times shown in Figure 3.22 (notice that the spikes are more random in the inhibitory cells, the top 50
cells of each plot). This appears somewhat at odds with claims that inhibitory interneurons play the primary role in synchronizing cortical neurons (e.g. Galarreta & Hestrin, 2001; Beierlein, Gibson, & Connors, 2000). However, this is almost certainly an artifact of using integrate-and-fire (IAF) model neurons as opposed to more detailed Hodgkin-Huxley style neurons - a substitution that was made solely out of concerns for computational speed in the simulations. Examination of the reduced phase models for IAF cells in these simulations indicated that the excitatory neurons were relatively synchronous in this firing rate range compared to many other Type I model neurons and that the inhibitory neurons were relatively asynchronous in this firing rate range compared to other Type I neurons (results not shown; see discussion in Section 2.2). If more detailed Hodgkin-Huxley neurons were used, the importance of inhibitory interneurons would likely increase. Future simulations will explore this possibility, and this issue is taken up further in the Conclusions and Future Directions (Section 5). However, the impact of the various factors we examine here will only be affected by this issue quantitatively rather than qualitatively.

Given this context, it is interesting to note that the effects of synaptic delays are somewhat different for synchrony among excitatory versus inhibitory cells. The addition of delays disrupts synchrony among the excitatory cells to a certain extent, whereas they can enhance synchrony among the inhibitory cells (at least for sufficiently long delays: compare delays of 0 and 2 ms in Figure 3.23 B). This result is somewhat expected since delays have the effect of making both excitatory and inhibitory interactions slower, where slow excitatory interactions disrupt synchrony and moderately slow inhibitory interactions promote synchrony (Gerstner, 1995, 1996; see also discussion in Section
2.2.1). The decrease in synchrony among the inhibitory cells from a delay of 0 to 1 ms is likely due to the attenuated synchrony in the excitatory cells rather than to an effect specifically on inhibitory interactions (note the large increase in synchrony among the inhibitory cells from a delay of 1 to 2 ms). Regardless of these differences, the results shown in Figure 3.23 A demonstrate that the addition of short synaptic delays do not radically undermine the results shown in Section 3.3.2, and indeed, under circumstances in which inhibitory interneurons contribute relatively more to the dynamics of activity, delays may in fact help to promote synchrony.

While local short synaptic delays of 1-2 milliseconds do not erase completely the repetition-related increases in synchrony, do longer-range interactions of cells with other cortical regions disrupt spike synchronization? In order to examine this, a second simple cortical region was allowed to interact with the first using inter-region delays of 5 ms and intra-region delays of 1 ms (similar to values used by Chawla et al., 1999). The second region was the same in all respects as the first (number of each type of cells, connection densities, etc.), except that it received its input from the first region rather than from the input neurons. Different amounts of feedback from the second region to the first were explored, ranging from 0% of the number of feedforward excitatory connections (i.e. feedforward connections between regions, and recurrent connections within a region) up to 50% in increments of 10%. The connections between regions were exclusively excitatory, and the feedback excitatory connections were more likely to synapse on excitatory rather than inhibitory cells (by a ratio of 2 to 1). These constraints were chosen to be broadly consistent with anatomical and physiological estimates of connection strengths/number between areas V1 and V2 measured in cortical slices (e.g.
Johnson & Burkhalter, 1996; Shao & Burkhalter, 1996). Changes in firing rate and coherence across stimulus repetitions are shown for the two interconnected cortical regions in Figure 3.24 for different amounts of feedback (region 2 => region 1: 0, 10, 30, and 50%). Error bars have been omitted to permit easier viewing. As in previous simulations, firing rates decreased in both regions across stimulus repetitions. Greater feedback from region 2 to region 1 led to slightly elevated firing rates in region 1. For lower levels of feedback, stimulus repetition led to increases in coherence in both regions. However, at higher levels of feedback, stimulus repetition only led to statistically significant coherence increases in region 1 (region 1: p<.0005 for 0-10,30-50% and p<.005 for 20%; region 2: p<.03 for 0%, p<.005 for 10%, p<.075 for 20-40%, p<.2 for 50%). With more runs of the simulation (i.e. >10), the small increases in coherence present in region 2 for higher levels of feedback would likely reach significance, although it is not clear if such small changes would have a substantial impact on subsequent processing (taken up in more detail in Simulation 4, Section 3.5).

Since neuromodulators tend to suppress excitatory feedback connections disproportionately in the cortex (e.g. Hasselmo & Cekic, 1996; Hasselmo, 1995), the actual amount of feedback present between different posterior cortical regions is unknown, although it is likely less than 50% - an estimate derived from slice physiology studies in the absence of neuromodulation (e.g. Shao & Burkhalter, 1996). The coherence values are also somewhat enhanced in region 2 relative to region 1. This may partially be due to the differences in firing rates, although it is also probably a result of each cell in region 2 having fewer independently firing inputs than cells in region 1.
Figure 3.24: Changes in firing rate and coherence for two inter-connected cortical regions separated by synaptic delays of 5 ms (within-region delays = 1 ms). Decreases in firing rate across stimulus repetitions are observed in both regions. Increases in coherence are also observed in both regions as long as the excitatory feedback from Region 2 to Region 1 is relatively weak (0-10%). Coherence increases in Region 1 are observed regardless of the amount of feedback.
In other words, there is more opportunity in region 2 for synchrony due to shared inputs by virtue of the convergence of inputs in the current model (1000 possible feed-forward inputs for region 1 cells, 250 possible feed-forward inputs for region 2 cells, and the same number of inputs is chosen for cells in both regions). Aside from this issue, region 1 inputs to region 2 fire less independently because their interactions help them to synchronize; the inputs to region 1 are independently firing Poisson processes chosen to minimize synchrony due to common inputs. Future simulations might alter this situation by including many more cells than per-cell connections in each cortical region. Taken together, the results of both of these simulations show that repetition-related changes in synchrony are not completely disrupted by synaptic delays and that such changes can propagate to subsequent cortical regions separated by longer delays, provided that the strength of positive feedback between regions is not too strong (see also Diesmann, Gewaltig, & Aertsen, 1999).

3.3.3 Impact of inter-stimulus interval on repetition-related changes in spike synchronization

In this simulation, the delay between two stimulus repetitions was varied in the model in order to examine the time course and persistence of the repetition-related coherence changes and to evaluate how well this time course matched that of short-term repetition priming effects observed with words in lexical decision and naming tasks. For example, McKone (1998) presented a series of words and non-words to adult human participants in a lexical decision task. Some of the words and non-words were repeated following a short delay ranging from 2 to 16 seconds and the benefit of stimulus repetition on lexical decision times was evaluated by comparing reaction times (RTs) to repeated
words/nonwords versus those that were only presented once (counterbalancing the words/nonwords that were repeated versus not repeated across participants). The time between repetitions was filled either with a blank delay or with intervening lexical decision trials. The magnitude of priming effects (RT difference between repeated and non-repeated stimuli) at different time delays is shown in Figure 3.25 A for words with and without intervening lexical decision trials. The benefit of repetition on reaction times both with and without intervening stimuli decreased substantially over the first 6-10 seconds, after which it approached relatively stable values (see McKone, 1995, for further discussion and review of short-term repetition priming effects). For comparison, changes in coherence have been plotted for the model across 2 stimulus repetitions separated by delays of .7, 1.9, 3.1, 4.3, and 5.5. seconds (Figure 3.25 B); these are coherence values for runs of the model that were discussed earlier in the context of addressing repetition suppression effects in neural recording experiments (see Section 3.2.2). As discussed in those previous simulations, stimulus durations were 500 ms and firing rates were calculated over the first 200 ms. Figure 3.25 C shows runs of the model reported above in the simulations of the fMR-Adaptation experiments (see Section 3.2.4); these simulations used a slightly longer stimulus duration (875 ms) and different values of delays (.125, 1.125, 3.125, and 7.125 seconds). It is apparent from these plots that the change in coherence across 2 stimulus repetitions in the model is strong for the first 1-2 seconds, after which it decreases substantially. The attenuation appears to be 4-5 seconds faster than for the human priming data. However, smaller significant increases in coherence are still present in the model over these durations (e.g. Figure 3.25B: \( p<.002 \) for 3.1 s, \( p<.025 \) for 5.5; Figure 3.25C: \( p<.004 \) for 3.125 s, \( p=.1 \) for 7.125 s).
Figure 3.25: The effect of inter-stimulus interval (ISI) on the magnitude of repetition priming effects in humans (A), as reported by McKone (1998), and the effect of ISI on changes in coherence observed in the model for simulations in Section 3.2.2 (B) and Section 3.2.4 (C). Repetition priming effects and coherence changes are reduced by longer delays between Stimulus 1 and 2, although this trend in the model is more rapid than for human priming data. From "The decay of short-term implicit memory: unpacking lag," by E. McKone, 1998, Memory & Cognition, 26, p.1181. Copyright 1998 by Psychonomic Society, Inc. Data re-graphed, permission pending.
Whether or not changes in coherence this small could produce significant changes in reaction times will be discussed further in Simulation 4 (see Section 3.5). The model clearly is unable to address priming effects at delays longer than 10 seconds. McKone (1998) has suggested that the priming effects present at delays longer than the first few seconds are due to longer-term learning mechanisms, largely because these priming effects are similar in magnitude to those observed at much longer delays of minutes. This may be the case, although such fast expression of long-term plasticity effects have yet to be observed neurophysiologically. If LTP/LTD or other more temporary NMDA-receptor-mediated plasticity effects can be expressed after durations of a few seconds, then they may indeed play some role in the current effects and would need to be included in the model. Future priming experiments that employ NMDA-receptor blockers would be critical for establishing this. Alternatively, slower firing-rate adaptation effects that have not been incorporated in the current model may help to extend the durations of the coherence changes, producing a better match to McKone's (1998) results. These possibilities will be considered further in the Conclusions and Future Directions (Chapter 5).

3.3.4 Comment

The current model demonstrates that short-term changes in spike synchrony accompany the decreases in firing rate that were observed in Simulation 1 (Section 3.2). As such, changes in synchrony may provide a principled explanation of short-term priming effects, at least those that span a few seconds, if the changes are sufficiently strong that they drive subsequent processing more effectively. Repetition-related changes in synchrony also
appear to be reasonably robust to input heterogeneity, local synaptic delays, and interactions with other cortical regions across longer delays - provided that excitatory feedback is not too strong. The assumption of weak feedback might be relaxed in future work if the inhibitory neurons in each region are parameterized to fire multiple spikes. This characteristic of some cortical networks has been shown to be critical for robust synchronization across longer synaptic delays (Ermentrout & Kopell, 1998; Karbowski & Kopell, 2000; Kopell, Ermentrout, Whittington, & Traub, 2000). The model produces weaker coherence changes across longer delays, consistent with weaker priming effects at longer delays. However, the durations are somewhat more transient than those reported for human priming data (McKone, 1998). This may reflect some degree of long-term plasticity that was not incorporated here or perhaps a slower short-term adaptation process (e.g. Sanchez-Vives et al., 2000; Schwindt et al., 1988). Along these lines, it is important to mention that a number of issues remain to be explored. These simulations, rather than representing the final word on the topic, are better viewed as a fruitful starting point to further understand the relationship between stimulus repetition and changes in firing rate and spike synchrony.

The current simulations indicate that firing-rate adaptation and synaptic depression help to synchronize excitatory cells following stimulus repetition by reducing their firing rates. Given the observations that increasing input heterogeneity, and hence firing-rate variance, while controlling for mean firing rate disrupts spike synchrony and that synaptic depression reduces the variance in addition to the mean firing rate across repetitions (Section 3.3.1), synaptic depression and firing rate adaptation may also aid spike synchronization by reducing the variability in firing rates. Without more direct and
rigorous analyses, it is hard to determine how much of the changes in synchrony are due to reducing the mean versus the variance in firing rates. However, it seems likely that both effects will contribute to some extent. Previous work has clearly demonstrated the importance of lower firing rates for synchrony in networks of excitatory Type I cells (e.g. Chow, 1998; Hansel et al., 1995), and Appendix B shows that this is also the case for networks of excitatory-inhibitory pairs. Future work will have to focus more specifically on the role of synaptic depression and firing-rate adaptation in promoting spike synchrony through reducing firing-rate heterogeneity.

Finally, it is worth pointing out that similar issues have been investigated by Chawla et al. (1999) in a model of the neocortex with somewhat contradictory results to those presented here. Chawla et al. (1999) found in a relatively detailed model of two interacting neocortical regions that phase-locked responses (referred to by them as synchronous responses) tended to occur for higher firing rates rather than lower firing rates. They also found that the incidence of phase-locked responses followed the reductions in the effective membrane time constants of the cells as they fired at higher rates (see Appendix D for mathematical explanation). Since the length of the membrane time constant of a cell determines how long it retains previous depolarizations, they argued that very short membrane time constants in their model served as a sort of filter for synchronous inputs such that cells could only spike when a substantial number of their inputs happened to align in time. Such effects may indeed occur in some circumstances, although there are a number of potential reasons that they did not observe synchronous spiking at lower rather than higher rates. The implementation of firing rate adaptation in the Chawla et al. (1999) model (actually a slow GABAb current for each of
the pyramidal cells onto itself) had such a slow rise time (τ = 30-90 ms) that it may have served to de-synchronize responses (synaptic depression was not included). They also included noisy background currents for each of the cells in their model, although they did not report the magnitudes of the instantaneous fluctuations; large independent noise provided to each cell would clearly reduce the ability of the model to synchronize. Lastly, if the magnitudes of the NMDA currents in their model were too large, they would also tend to desynchronize responses due to the slow excitatory kinetics of those currents. Any of these circumstances could in principle de-stabilize synchronous spiking at lower firing rates. Indeed, under conditions that prevent any stable phase-locked solutions, the only times that spikes might occur near-synchronously at above-chance levels would be at higher firing rates with very short effective membrane time constants (e.g. < 1 ms). However, if this model were to incorporate synaptic depression, it - like standard firing rate models - would appear to predict slower rather than faster responses following stimulus repetition since both rate and synchrony would be decreasing (see Appendix A for further discussion). Future work remains to be done to understand better the discrepancies between the results of the two models. For the time being, the current model provides useful insights into how synchronous spiking might occur at lower rather than higher firing rates, and thus, it is worth exploring further as an account of the relationship between short-term repetition suppression and short-term behavioral priming.
3.4 Simulation 3: Contribution of Inhibition, Adaptation, and Synaptic Depression to Synchronization

The previous simulation demonstrated that short-term firing rate decreases due to firing-rate adaptation and synaptic depression were also accompanied by enhanced spike synchronization. This simulation evaluated the contribution of these two short-term plasticity mechanisms, as well as the contribution of inhibition, to spike synchronization by turning off one or more mechanisms. Additional dynamical effects (i.e. effects on the phase response curves, PRCs) of these mechanisms beyond that of reducing the means and variances of firing rates were evaluated by blocking each mechanism and controlling for differences in firing rate across conditions (see Section 2.2 for further discussion). All three mechanisms were predicted to aid synchrony by reducing rates, but only inhibition and adaptation were expected to aid synchrony further by altering the PRCs of the excitatory cells (e.g. Crook et al., 1998; Ermentrout et al., 2001; Ermentrout & Kopell, 1998; see also Appendices B and C).

3.4.1 Effect of blocking inhibition, adaptation, and synaptic depression on firing rate and synchrony

The first step in understanding these issues was to block inhibition ($I_{GABA_a}$), firing-rate adaptation ($I_{K(Ca)}$), and synaptic depression one at a time and measure the resulting changes in firing rate and coherence for the excitatory cells. As in previous simulations, stimulus durations were 500 ms, firing rate was calculated over the first 200 ms, and coherence values included spikes that occurred between 20 ms post-stimulus onset and stimulus offset. Stimuli were presented 10 times in each condition separated by an ISI of
1 second, and conditions were repeated 10 times to yield stable estimates of firing rate and coherence. Figure 3.26 A shows the effect of blocking each mechanism on the mean and standard deviation of firing rates relative to the normal network with +/- 20% input heterogeneity. The effect of blocking synaptic depression on firing rates is not shown because the rates lie far outside the range of the other conditions [mean firing rate = 396.52 +/- 28.28 Hz (SD)]. The increases in firing rate relative to those in the normal network were least severe for the blocking of inhibition (\(I_{\text{GABA}_A}\)), they were more severe for the blocking of firing-rate adaptation (\(I_{K(Ca)}\)), and they were the most severe for the blocking of synaptic depression (raising firing rates in the excitatory cells to levels near the upper bound possible, 500 Hz, for neurons with refractory periods of 2 ms). The standard deviations of firing rates were also substantially increased for blocked adaptation and synaptic depression.
Figure 3.26: The effect of blocking inhibition ($I_{GABA}$), firing-rate adaptation ($I_{K(Ca)}$), and synaptic depression on firing rates (A) and coherence values (B). Blocking all three mechanisms led to increases in the means and standard deviations of firing rate and decreases in coherence.
Changes in coherence values for these conditions are what one would expect for firing rates with increased means and standard deviations: The condition with the least firing rate increases, blocked inhibition, showed mild decreases in coherence relative to the normal network. The other two conditions showed larger coherence decreases with no repetition-related changes. These results are consistent with the notion that inhibition, adaptation, and synaptic depression all help to enhance spike synchronization by dynamically decreasing firing rates.

3.4.2 Contributions of inhibition, adaptation, and synaptic depression to synchrony that are unrelated to changes in firing rate

This simulation evaluated potential dynamical contributions of inhibition, adaptation, and synaptic depression to changes in spike synchronization beyond that achieved by reducing firing rates. This was done by blocking each mechanism individually and roughly matching firing rates across conditions by scaling down the excitatory synaptic strengths in the model. In order to match both the means and standard deviations of firing rate in each "blocked" condition with those of the normal network, it was necessary to choose runs of the model with the appropriate level of input heterogeneity (prior to blocking), so that after blocking the particular mechanism and scaling down the excitatory synaptic strengths, similar steady-state firing rates would result. For example, the impact of firing-rate adaptation ($I_{K(Ca)}$) on synchrony was evaluated by setting the adaptation current to 0 in the normal network with +/- 0% input heterogeneity and scaling down the excitatory synaptic strengths by 0.725. This resulted in an initial firing rate of 32.6 +/- 6.7 (SD) Hz to the first stimulus and a final firing rate of 13.9 +/- 6.4 (SD) Hz to the tenth stimulus. The asymptotic firing rates following stimulus repetition (mean,
standard deviation) were comparable to those in the normal network with +/- 40% heterogeneity (10th stimulus: 12.0 +/- 5.5 Hz). Inhibition and synaptic depression were similarly blocked and compared to the normal network with +/- 40% heterogeneity. The values of coherence for each of the conditions are shown across repetitions in Figure 3.27. Consistent with prior work (e.g. Crook et al., 1998; Ermentrout et al., 2001), blocking firing-rate adaptation reduced changes in coherence across repetitions. Somewhat unexpected, however, were the effects of blocking inhibition and synaptic depression. Blocking inhibition ($I_{GABA_a}$) actually increased repetition-related changes in coherence once firing rates were controlled. The apparent de-synchronizing role of inhibition in this model, as discussed previously in the section on synaptic delays (Section 3.3.2) is likely an artifact of employing the IAF model neurons as an approximation to Hodgkin-Huxley style neurons (see also Figures 3.22 and 3.23). Future simulations with more detailed Hodgkin-Huxley neurons are expected to reveal a more synchronizing role for inhibition in these firing-rate ranges (e.g. Ermentrout et al., 2001; Ermentrout & Kopell, 1998). The results for synaptic depression are also interesting. As in the case of adaptation, blocking synaptic depression and matching steady-state firing rates reduces coherence values. This might imply that synaptic depression, like adaptation, alters the dynamics of the excitatory-inhibitory interactions to promote synchrony beyond any simple effects of reducing firing rates. However, this is somewhat inconsistent with the results of the phase model analyses for synaptic depression shown in Appendix C.
Figure 3.27: The effect of blocking inhibition ($I_{\text{GABA}}$), firing-rate adaptation ($I_{K(Ca)}$), and synaptic depression on coherence values when asymptotic firing rates (mean and standard deviation) were matched across conditions. Blocking adaptation or synaptic depression reduced coherence values, whereas blocking inhibition actually increased them.

One possible resolution of these two findings is that synaptic depression and firing-rate adaptation interact. Ermentrout et al. (2001) have shown that progressively increasing the magnitude of firing-rate adaptation while keeping rates constant can lead phase-locked solutions to move closer to synchrony. Since the magnitude of adaptation depends on firing rate, starting out at firing rates that are too low in the model may not allow adaptation to build up sufficiently to play a synchronizing role. This appears to be the case for the model: When the strengths of the excitatory synapses in the normal model (+/- 0% heterogeneity) are scaled down to yield lower rates (e.g. scaling down by 0.6: 1st stimulus yields 13.0 +/- 3.0 Hz, 10th stimulus yields 6.8 +/- 2.5 Hz), coherence values are decreased to the same levels as when adaptation is blocked (~ 0.17 across
stimuli). Similarly, when synaptic depression was blocked, the excitatory synaptic strengths had to be scaled down dramatically to match the steady-state firing rates of the normal model, yielding low initial firing rates that remained relatively constant across repetitions (8.5 +/- 3.3 Hz) and preventing strong adaptation. If firing rates could start out initially large and then rapidly decrease during the processing of a stimulus into a lower range that permits better synchrony, adaptation might be strong enough to play a synchronizing role. This is precisely what the faster term of synaptic depression ($\tau$ of $\sim$ 600 ms) helps to accomplish. Future work will be aimed at clarifying the interaction between adaptation and synaptic depression in more detail.

3.4.3 Comment

As expected, when inhibition, adaptation, and synaptic depression are blocked, firing rates increase and synchronization breaks down. When the means and variances of firing rates were matched with that of the full model, adaptation and synaptic depression appeared particularly critical for allowing neurons to synchronize across repetitions. Even though the adaptation used here was not slow enough to accrue across repetitions, it had an important impact on neuronal dynamics during each stimulus repetition, interacting with synaptic depression as it built up across repetitions. Blocking either process substantially attenuated synchronization. When taken together with the results of Appendix C and the impact of scaling down excitatory synaptic strengths in the full model, the results with blocked synaptic depression may have more to do with the lack of strong adaptation that builds up with higher initial rates. This suggests that adaptation and synaptic depression together play a crucial role in altering the dynamics of excitatory
interactions. If Hodgkin-Huxley style neurons are used, inhibition may be found to play a similar role to that of adaptation for the range of firing rates explored here (e.g. Ermentrout et al., 2001).

3.5 Simulation 4: Effect of Firing Rate and Synchronization on Reaction Time

Here we explore ways in which changes in firing rate and spike synchronization in posterior cortical regions might influence the generation of motor responses in anterior cortical regions. As discussed above, stimulus repetition most often induces firing rate decreases in posterior cortical regions that are involved in perceptual processing. The argument was made in Simulation 1 above (Section 3.2) that the activity decreases observed over the relatively short time scale of seconds are due to the buildup of short-term plasticity mechanisms such as firing-rate adaptation and synaptic depression. However, activity decreases are not observed in all cortical regions in neuroimaging studies in all behavioral contexts. For example, Jiang et al. (2000) showed that repetition of distractor stimuli over several seconds in a delay match-to-sample task had little or no effect on activity levels in more anterior cortical regions (e.g. frontal/insular, primary, and supplementary motor cortex), whereas they led to decreased activity in posterior regions such as ventral temporal and parietal cortex. van Turennout et al. (2000) showed than short-term repetition of picture stimuli in silent naming led to significant activity increases in left frontal/insular cortex while simultaneously leading to decreases in
occipital and temporal cortical regions. It is possible that the lack of short-term activity decreases in these cases is due to some differences in cellular properties in different brain regions, where adaptation and/or synaptic depression are weaker or absent in more frontal regions. It is also possible that the lack of decreases is due to differences in neuromodulation in the different regions. A variety of neuromodulatory factors alter transmitter release (potentially affecting synaptic depression) and suppress the potassium currents mediating adaptation effects (see Hasselmo, 1995, for a review). Chapter 4 will take these issues up in more detail for the neuromodulatory actions of acetylcholine and norepinephrine (see also Section 1.1). The differences in repetition-related activity changes in frontal versus more posterior cortical regions may indicate that these regions play somewhat different roles in various tasks. Prefrontal cortex has been argued to be important for a wide range of higher level cognitive functions such as attention, working memory, cognitive control, problem solving, and reasoning (e.g. Cohen & Servan-Schreiber, 1992; Miller & Cohen, 2001; Shallice, 1988; Stuss & Knight, 2002). Sustained prefrontal activity may help to maintain task context, behavioral goals, and aid in the integration and conjunction of sensory events/information occurring over time. Indeed, strong effects of firing-rate adaptation and synaptic depression might disrupt such functions if reductions in activity were too great. The strong synchronous dynamics induced by adaptation and synaptic depression in Simulations 2 and 3 above (Sections 3.3 and 3.4) might also disrupt sustained activity and slow temporal integration because transient excitatory synaptic currents occurring at the same time can decay away completely in between synchronous spike volleys (e.g. Compte, Brunel, Goldman-Rakic, & Wang, 2000; Gutkin, Laing, Colby, Chow, & Ermentrout, 2001). Accordingly, the
current simulation explores how dynamic changes in firing rate and synchrony in posterior cortical regions may impact the activity in more frontal/motor regions that are assumed to exhibit weaker short-term plasticity effects.

How then might decreases in neural firing rate and increases in spike synchronization throughout posterior cortical regions lead to the earlier initiation of a motor response in anterior cortical regions? As discussed earlier in Section 1.4.3, Hanes and Schall (1996; Schall, 2001) demonstrated that response time in an eye movement task is well predicted by the spiking activity of individual motor cortex neurons. In particular, reaction times are determined by the time it takes neural firing rates to build up to a critical fixed threshold value. A full simulation of how changes in input firing rate and spike synchrony might influence motor responses would involve implementation of an entire population of prefrontal/motor cortex neurons, as well as the posterior cortical regions serving as input. The current simulation develops a highly simplified approximation of this larger situation: Firing rate and spike synchrony were manipulated orthogonally in a population of input neurons, 200 excitatory and 50 inhibitory, and the resulting firing rates in a single output "motor" neuron were recorded. This is admittedly a less developed approach than some of the previous simulations, and more complete simulations are planned for future work (see Section 5.2). Nevertheless, to the extent that the firing rate of a single output neuron averaged across many trials behaves similarly to the average firing rate of a larger population of cells when input firing rates and synchrony are increased (e.g. output rate increases monotonically), the results obtained should not be unduly misleading. Each input neuron fired with Poisson characteristics at a specified rate (ranging from 0-50 Hz in increments of 5 Hz), and the
degree of input synchrony was manipulated by assigning a certain number of independent Poisson processes to the 250 inputs. Greater synchrony was achieved by reducing the number of independent Poisson processes (e.g. assigning 1 process to all inputs corresponds to perfect input synchrony), and weaker synchrony was achieved by increasing the number of processes (e.g. assigning 250 independent processes to the 250 input cells, 1 to a cell, leads to highly asynchronous firing). Half of the excitatory input neurons (100) were intended to represent spiking activity in posterior cortical regions, and the remaining excitatory and inhibitory inputs were intended to represent spiking activity in neighboring motor cortex cells. In some conditions, the firing rates of the neighboring motor inputs followed the rates of the posterior inputs while the times of the spikes were relatively asynchronous (each input getting its own independent Poisson process). In other conditions the motor inputs followed both the firing rate (0-50 Hz) and degree of synchrony present in the posterior inputs (assigned between 1 and 10 independent Poisson processes, with each group of inputs drawing on its own unique pool of processes). These two cases represent two different extremes: 1) motor cortex neurons fire asynchronously, irrespective of changes in synchrony of neurons in posterior cortex, and 2) motor cortex neurons reflect completely the degree of synchrony present in their inputs. As just mentioned above, too much synchrony present in frontal cortical regions might disrupt the ability of these regions to slowly integrate their inputs and gradually build firing rates up to the threshold for initiating a response.

Response times were estimated in this simple model by measuring the amount of time it took the single output motor neuron to reach a criterion number of spikes. While this measure is not precisely the same as the amount of time required for firing rate in
individual motor cortex cells to build up to a fixed firing rate threshold, the two measures will tend to be highly correlated. This is true because firing rate activity that builds more rapidly toward a fixed threshold reflects faster instantaneous firing rates, and faster instantaneous rates will in turn lead a criterion number of spikes to be reached earlier (see Figure 1.4). The implementation and parameters of the single output neuron were as described in Section 3.1 with two main exceptions: 1) firing-rate adaptation and synaptic depression were either not present ($g_{K(Ca)}$ and $d^*$ set to 0) or were included only weakly to examine their impact on the results ($g_{K(Ca)}$ and $d^*$ reduced by 80%), and 2) the constant background currents used in previous simulations were replaced here with stochastic background EPSPs and IPSPs that depolarized the resting membrane potential ($V$) by 8-10 millivolts (mV) and led to fluctuations in $V$ with a standard deviation of 3-4 mV and spontaneous firing rates of $\sim$ 1 Hz. These values were chosen to be roughly consistent with the in vivo neural recording data of Destexhe and colleagues (e.g. Destexhe & Pare, 1999; Destexhe, Rudolph, Fellous, & Sejnowski, 2001; Pare, Shink, Gaudreau, Destexhe, & Lang, 1998). The inclusion of stochastic background synaptic currents introduced an additional degree of randomness in the output neuron's responses, and it helped to evaluate how changes in input synchrony might influence responses under reasonable assumptions about the average membrane time constant of the cell (constant leak conductance plus the background currents yielded a time constant of $\sim$ 20 ms; see Appendix D for further discussion).
3.5.1 The Impact of Input Firing Rate, Input Synchrony, and Balanced Excitation/Inhibition on Output Firing Rate

Salinas and Sejnowski (2000) have previously examined the impact of input firing rate and the degree of synchrony on the output firing rate of a simple integrate-and-fire (IAF) model cell. They found that as long as the balance of excitatory and inhibitory synaptic currents was relatively even (i.e. little or no net change in the average membrane potential), increases in input synchrony of either the excitatory or inhibitory cells could lead to higher firing rates in the output cell. Indeed, synchrony in this case appeared to have a multiplicative impact on output rates, leading them to argue that changes in input synchrony can be tantamount to changes in synaptic gain. This is somewhat understandable when one considers that it is primarily fluctuations in the synaptic current that drive the output cell to fire when the average membrane potential is sufficiently below spiking threshold: Synchrony among either excitatory or inhibitory cells enhances fluctuations because spikes occurring together will lead to large transient increases and decreases in current, causing more output spikes. In contrast, if the balance of excitation to inhibition favored excitation too much or if synchrony between the excitatory and inhibitory cells increased, output responses were primarily determined by input firing rates. When excitatory and inhibitory neurons fire at the same time, the fluctuations in the synaptic current do not increase much because the depolarizing and hyperpolarizing currents tend to cancel each other out. However, synchrony can still have a significant impact in these cases, depending on the heights of the individual synaptic currents, the number of input cells, and their firing rates; there will be a greater impact of synchrony on output firing rate for smaller synaptic currents, fewer inputs, and lower firing rates. The work of Salinas and Sejnowski (2000) therefore established that the balance of
excitation to inhibition in cortical cells is an important determinant of whether increases in input synchrony can lead to heightened output responses.

In the current simulation, the impact of input firing rate, the degree of input synchrony, and the balance of excitation/inhibition on the firing rate of the output neuron was investigated. As mentioned above, input firing rates ranged between 0 and 50 Hz in increments of 5 Hz. Input synchrony was varied orthogonally to rate at 10 different levels, assigning from 1 to 10 independent Poisson processes to the inputs in increments of 1. This corresponded to average coherence values of 1.0 for 1 independent Poisson process down to 0.16 for 10 processes. These particular values of rate and coherence were chosen to span the values observed in Simulations 1 and 2. For conditions in which neighboring motor cortex inputs increased in synchrony along with the posterior cortical inputs, the excitatory and inhibitory neurons were assigned inputs from their own separately generated pools of Poisson processes. This was intended to minimize the covariation of excitatory-inhibitory neuron synchrony with excitatory-excitatory and inhibitory-inhibitory neuron synchrony (Salinas & Sejnowski, 2000), and it was intended to capture the lack of cross-group synchrony that might occur due to inter-region synaptic delays. The balance between excitatory and inhibitory synaptic currents in the model was manipulated by scaling up or down the value of the inhibitory synaptic conductance while leaving the excitatory synaptic conductance fixed; three different values of $g_{EI}$ were explored (0.15, 0.225, and 0.3; $g_{EE}=0.02$). Figure 3.28A shows the impact of these various factors on the firing rate of the single output cell with firing-rate adaptation and synaptic depression blocked ($g_{K(Ca)}$ and $d^*$ set to 0) and changes in synchrony only occurring for the posterior cortical inputs. Figure 3.28B shows the same conditions under
weak adaptation and synaptic depression ($g_{K(Ca)}$ and $\alpha$ scaled down by 0.2). With adaptation and synaptic depression blocked, the left side of Figure 3.28A shows that increases in input firing rate lead to increases in the firing rate of the output cell ($p<.0005$ for all conditions), although these changes are larger when the balance of excitation to inhibition favors excitation (e.g. compare $g_{EI}=0.15$ with the approximately balanced case of $g_{EI}=0.3$). The right side of Figure 3.28A shows that increases in coherence lead to increases in output firing rate only for relatively balanced excitation/inhibition (e.g. $g_{EI}=0.3$, $p<.0005$), consistent with the results of Salinas and Sejnowski (2000). If excitation is relatively strong, output firing rate actually decreases when coherence is increased (e.g. $g_{EI}=0.15$, $p<.0005$). This can be understood if one thinks about the extreme case of each EPSP being strong enough to cause a spike by itself: Having the spikes align will actually reduce the number of spikes possible because of the refractory period of the output cell and the relatively fast decay of $I_{AMP}$. Scaling the heights of the EPSPs to small enough values will eventually lead synchrony to be favored again because it will require several cells firing at the same time to reach threshold. It is also important to note that for the balanced excitation/inhibition case, most of the increases in output firing rate occur over relatively modest increases in coherence (e.g. 0.2-0.4). The left panel of Figure 3.28B shows that with weak adaptation and synaptic depression included, increases in input firing rate yield only limited increases in output rate. As input rates grow large, the output firing rate decreases again, revealing a non-monotonic relationship between input and output rates. This occurs because synaptic depression is larger at excitatory relative to inhibitory synapses (see Figure 3.4 in Section 3.1.2), and this difference becomes more and more emphasized as input rates increase.
Figure 3.28: The impact of changes in input firing rate, input coherence (for the posterior cortical inputs only), and balanced excitatory/inhibitory synaptic currents on the firing rate of a single output cell that represents motor cortical responses. (A) shows the results for different levels of excitatory/inhibitory balance with adaptation and synaptic depression completely blocked (evenly balanced: $g_{EI}=0.3$, stronger excitation: $g_{EI}=0.15$), whereas (B) shows the same results for 20% adaptation and synaptic depression. See text for explanation.
Effectively, synaptic depression is leading to a shift in the balance of excitation/inhibition to favor inhibition; this is precisely what Varela et al. (1999) proposed following their observation of this asymmetry. As the balance of excitation/inhibition shifts dynamically during a stimulus to favor inhibition, the output cell becomes more sensitive to input synchrony. The right panel of Figure 3.28B shows that this is indeed the case: Increases in coherence lead to enhanced output firing rate for all three values of $g_{EI}$ ($p<.0005$ for all conditions). Synaptic depression, to the extent that it is present, may therefore play a role in helping neurons in anterior cortical regions respond more vigorously to increases in the synchrony of neurons in posterior cortical regions.

Similar results were found when changes in synchrony were applied to the neighboring motor cortical inputs, as well as to the posterior cortical inputs. Figure 3.29A shows the impact of input firing rate, input synchrony, and the balance of excitation/inhibition on output firing rate with adaptation and synaptic depression blocked ($g_{K(Ca)}$ and $d^*$ set to 0). Figure 3.29B shows the same conditions under weak adaptation and synaptic depression ($g_{K(Ca)}$ and $d^*$ scaled down by 0.2). The left panel of Figure 3.29A shows that output firing rate increased with increases in input rates for all of the excitation/inhibition conditions ($p<.0005$ for all conditions). The right panel shows that increases in output rate were steeper over modest increases in input coherence and decreased slightly for larger coherence values (compare to Figure 3.28A). When weak adaptation and synaptic depression were included, the relationship between input and output firing rates again became non-monotonic, and the output cell became more sensitive to changes in input coherence (Figure 3.29B).
Figure 3.29: The impact of changes in input firing rate, input coherence (for all inputs), and balanced excitatory/inhibitory synaptic currents on the firing rate of a single output cell that represents motor cortical responses. (A) shows the results for different levels of excitatory/inhibitory balance with adaptation and synaptic depression completely blocked (evenly balanced: $g_{EI}=0.3$, stronger excitation: $g_{EF}=0.15$), whereas (B) shows the same results for 20% adaptation and synaptic depression. See text for explanation.
If anything, allowing neighboring motor cortical inputs to synchronize along with the posterior cortical inputs increases the sensitivity of the output neuron to changes in synchrony, at least over the range of moderate coherence values.

### 3.5.2 The Impact of Input Firing Rate and Synchrony on Reaction Time

As discussed above, there is some reason to believe that neurons in more anterior cortical regions fire relatively asynchronously as they integrate information present in their inputs and exhibit little or no repetition-related decreases in firing rate. We will therefore focus in more detail on the case in which the neighboring motor cortical inputs fired relatively asynchronously and for which firing rate adaptation and synaptic depression were blocked (see Figure 3.28A). If response times improve in this case as the firing rates of inputs decrease and the spike times of the posterior cortical neurons synchronize then they will surely do so for cases in which the motor cortical inputs also synchronize and synaptic depression is incorporated. It is clear from the results depicted in Figure 3.28A that changes in synchrony will only increase output firing rate and decrease response times if the balance between excitation and inhibition is relatively even, such as when $g_{EI}=0.3$. Figure 3.30 shows reaction times for this condition as a function of input firing rate and the degree of synchrony present in the posterior cortical inputs. Reaction times were estimated as the latency required to reach the 10th output neuron spike following the onset of input spiking. Each combination of input firing rate and coherence was repeated 20 times to yield a relatively stable estimate of reaction time, and the median of these 20 values was calculated. Figure 3.30A shows the impact of input firing rate on reaction time, averaged across the medians in the different coherence conditions.
Figure 3.30: The impact of changes in input firing rate and input coherence (in the posterior cortical inputs only) on reaction time for evenly balanced excitatory/inhibitory synaptic currents and no adaptation or synaptic depression. (A) shows that increases in input firing rate generally lead to faster reaction times, as do increases in input coherence, shown in (B). (C) shows that the coherence-related changes in reaction times happen mainly in the higher range of input firing rates.
Increases in input firing rate from 5 to 25 Hz led to substantial decreases in reaction time, with slight increases in reaction time for larger rates (the reverse pattern is also present in output firing rate in Figure 3.28A). The decreases in response time occur because increases in input firing rate for balanced excitation/inhibition also give rise to modest increases in the fluctuations of synaptic currents. The reaction times increase again slightly for larger input rates likely because high-rate inputs will shorten the average membrane time constant of the output cell, allowing it to respond more rapidly to changes in the synaptic current but also weakening the absolute voltage changes (see Appendix D for an explanation of how changes in membrane currents alter cellular integrative properties). Reaction times are also decreased substantially by modest increases in input coherence over the range of 0.2 to 0.4, as revealed by Figure 3.30B. Figure 3.30C shows that most of these reaction time changes due to coherence occur over input firing rates larger than 15-20 Hz. For low enough input rates, changes in coherence have very little impact and only changes in firing rate matter. It is also clear from this figure that even with 20 runs of each condition, median reaction times are still relatively noisy. Nevertheless, an ANOVA calculated over the individual reaction times revealed highly significant main effects of input rate [$F(9, 1900)=17.24, p<.0005$] and coherence [$F(9, 1900)=49.42, p<.0005$], as well as a highly significant interaction indicating larger reaction time decreases for coherence increases at larger input firing rates [$F(81, 1900)=2.01, p<.0005$].
3.5.3 Estimating Reaction Times for Simulation 2 (Section 3.3)

Using the results of the last section that related input firing rate and degree of spike synchronization to reaction times under the assumption of relatively balanced excitation/inhibition (see Figure 3.30), responses were estimated for runs of the model presented in Simulation 2 (Section 3.3) in order to evaluate if the changes in coherence observed were sufficient to produce repetition priming effects. This was possible because the values of input firing rate and coherence used in the section above spanned the entire range of values measured in Simulation 2. Reaction times were calculated for each new combination of rate and coherence by interpolating between known values that corresponded to adjacent rate and coherence values. Adjacent values were simply weighted as an inverse function of their distance (1/Distance) from the new point in the rate X coherence map and then averaged, normalizing by the sums of the inverse distances. Using this simple method, reaction times were estimated for the results previously displayed in Figure 3.21 (Section 3.3.1) regarding the impact of stimulus repetition and input heterogeneity on firing rate and coherence. Recall that four different levels of input heterogeneity were explored in that simulation: +/- 0, 20, 40, and 60% of the number of excitatory inputs. The results are shown below in Figure 3.31. It is clear from this figure that strong reaction time decreases occur from Stimulus 1 to Stimulus 2 for all four levels of input heterogeneity ($p<.0005$ for all conditions), reflecting highly significant repetition priming effects. This is particularly interesting given that the coherence increases for larger values of input heterogeneity were only around 0.1 (e.g. +/- 40 and 60%). Significant priming effects are also observed for later stimulus repetitions when responses are compared relative to the first stimulus. However, reaction
times do increase past the second or third stimulus repetition. This occurs in the model because the increases in coherence that build for later stimuli are not large enough to offset the continued decreases in firing rate. Recall also from Figure 3.30C above that changes in coherence do not have a large impact on reaction times at lower firing rates, whereas changes in rate do. This pattern of results is interesting with regard to behavioral phenomena such as semantic satiation, where significant priming effects are observed after a small number of prime exposures that attenuate with a larger number of prime exposures (e.g. Balota & Black, 1997; Smith, 1984). This point will be taken up further in Chapter 4 (particularly in Simulation 8, Section 4.5), as well as in the later section on Future Directions (Section 5.2).

3.5.4 Comment

The above simulations demonstrated that as long as excitatory and inhibitory synaptic currents are relatively balanced, changes in the spike synchronization of neurons in posterior cortical regions may be capable of eliciting larger firing rates in more anterior cortical regions, thus evoking earlier motor responses. Relatively weak effects of synaptic depression and heightened synchrony in neighboring motor cortical cells appeared capable of leading to priming effects even if the assumption of balanced excitation/inhibition were relaxed to a certain extent. While the actual status of balanced excitation/inhibition in different cortical regions is unknown, a number of theorists have argued that it is likely to be more balanced that previously expected (e.g. Miller & Troyer, 1997; Salinas & Sejnowski, 2000; Shadlen & Newsome, 1994).
Figure 3.31: The impact of stimulus repetition and input heterogeneity on reaction time for the simulations presented in Section 3.3.1. Reaction times are decreased strongly for all heterogeneity conditions from Stimulus 1 to Stimulus 2, showing significant repetition priming effects. Reaction times increase slightly in all conditions for later stimulus repetitions, although they are generally still faster than the response to the first stimulus.

The heights of EPSPs and IPSPs in the model (1-4 mV), as well as the ranges of firing rates for excitatory and inhibitory cells (0-100 Hz), are well within the range of those observed in real cortical neurons. Further evaluation of this issue will have to await additional empirical work in which the relative balance of excitatory/inhibitory synaptic currents and changes in this balance due to synaptic depression are estimated for different cortical regions - particularly those directly involved in generating motor responses. As discussed earlier, additional simulation work is also needed to explore a more full implementation of the mechanisms involved in response generation in motor cortical
regions so that the results observed by Hanes and Schall (1996) can be addressed more directly (see Section 5.2 for elaboration of this point).

3.6 Simulation 5: Quantification of Efficiency

In this simulation, it will be demonstrated that the changes in firing rate observed in Simulation 1 (Section 3.2) and the estimated changes in response time characterized in Simulation 4 (Section 3.5) lead to enhanced processing efficiency. An appropriate efficiency (inefficiency) measure should increase (decrease) if energy expenditure is reduced with performance remaining constant, and it should also increase (decrease) if performance improves without changes in energy use. Accordingly, two different measures of efficiency were explored. The first was calculated quite simply by multiplying the average firing rate for each stimulus by its corresponding response time estimated in Simulation 4. This essentially gives a measure of the number of spikes per cell that are required to elicit a response, and as such it is a rough measure of metabolic inefficiency. Decreases in this measure will be taken as improved metabolic efficiency: Improved performance (reduced reaction times) with fixed energy expenditure or reduced energy expenditure (lower firing rates) with fixed performance will correctly lead to decreases. Results are shown in Figure 3.32 for the four different levels of input heterogeneity investigated in Simulations 2-4.
Metabolic inefficiency was significantly reduced for each heterogeneity condition across stimulus repetitions (all conditions: $p<.0005$), with little or no differences among conditions. Perhaps the only surprising aspect of these results is that inefficiency continues to be reduced even for later stimulus repetitions as reaction times climb slightly from repetitions 2 through 10 (see Figure 3.31). A more general, scale-free metric of metabolic efficiency was also designed:
\[ \Delta \text{Efficiency} = \frac{1}{2} \left( \frac{e_1 - e_2 + p_2 - p_1}{e_1 + e_2 + p_2 + p_1} \right) \]

where \( e \) and \( p \) refer to energy use and performance, respectively, for Stimulus 1 and 2. Here changes in \( e \) and \( p \) contribute separately to changes in efficiency, and normalizing each term by the sum of its individual values guarantees that proportional decreases in energy use and proportional increases in performance from Stimulus 1 to Stimulus 2 will contribute equally to increases in efficiency. Since firing rate \( (f) \) is directly related to energy use and reaction time \( (RT) \) is inversely related to performance, we have:

\[ \Delta \text{Efficiency} = \frac{1}{2} \left( \frac{f_1 - f_2}{f_1 + f_2} + \frac{1}{RT_2} - \frac{1}{RT_1} \right) \]

which simplifies to

\[ \Delta \text{Efficiency} = \frac{f_1 \cdot RT_1 - f_2 \cdot RT_2}{(f_1 + f_2) \cdot (RT_1 + RT_2)} \]

Positive values of this metric indicate that efficiency has improved from Stimulus 1 to Stimulus 2 (maximum of +1), whereas negative values indicate that efficiency has decreased (minimum of -1). Figure 3.33 re-graphs the results shown above for the four different levels of input heterogeneity using this new measure. Changes shown for each stimulus repetition were calculated relative to Stimulus 1 firing rate and reaction time.
Metabolic efficiency increased significantly across stimulus repetitions for each heterogeneity condition (all conditions: \( p < .0005 \)), with little or no differences among conditions. As with the simpler metric, efficiency continued to improve even as reaction times climbed slightly following the second stimulus repetition.

3.6.1 Comment

As expected from the results of Simulations 1-4, metabolic efficiency was shown to improve across stimulus repetitions for two different measures of efficiency. Efficiency
continued to improve even when the reaction times estimated in Simulation 4 increased slightly because the proportional decreases in firing rate were larger. This demonstrates that, at least for this model, the short-term plasticity processes of firing-rate adaptation and synaptic depression help to improve processing efficiency in the short term, yielding better performance for less energy.