

MECHANISMS UNDERLYING ENHANCED PROCESSING EFFICIENCY IN
NEURAL SYSTEMS

by

STEPHEN JAMES GOTTS

A Dissertation Presented to the
FACULTY OF THE DEPARTMENT OF PSYCHOLOGY
CARNEGIE MELLON UNIVERSITY

In Partial Fulfillment of the
Requirements for the Degree
DOCTOR OF PHILOSOPHY
(Psychology)

May 2003

Dedication

This thesis is dedicated to Dr. Frederick J. Bremner, who first ignited my curiosity about the relationship between brain and mind.

Acknowledgements

On leaving my home for the past six years, I feel very fortunate to have been surrounded by so many exceptional, gifted, and giving teachers. First and foremost among these, I would like to thank the members of my dissertation committee: David Plaut was the model advisor, giving me just the right amount of guidance, feedback, and space to let me develop my own ideas; aside from teaching me basic scientific and critical thinking skills, he imparted to me an invaluable meta-perspective on how one goes about developing an explanation of something. Carson Chow, in teaching me the basics of biophysics and computational neuroscience on which much of the current work is based, really acted as a second advisor - exposing me to the valuable perspectives of physics and mathematics. Jay McClelland, in addition to his always on-point gut reactions, provided an ideal example of the excellence and humility that all scientists should strive to achieve. Bobby Klatzky kept me honest by pushing me to think hard about the empirical implications of my ideas and encouraging me to pursue them, something that I'm just starting to do now. I also owe a deep debt of gratitude to Lisa Cipolotti and Tim Shallice, who graciously allowed me to spend a summer at the National Hospital in London so that I could stand over their shoulders and learn first-hand how Neuropsychology is done.

Much of my graduate education was the result of more informal discussions and interactions with friends and colleagues in the PDP research group, the CMU Psychology department, and the CNBC. One couldn't ask for a more vibrant and engaging research community. Among these colleagues, I'd particularly like to thank Marlene Behrmann, Brad Best, Bard Ermentrout, Lisa Gershkoff-Stowe, Mike Harm, David Noelle, Alex Petrov, Lynne Reder, Tim Rogers, Doug Rohde, and David Wood for pushing my mind in many different healthy directions.

No one can get through graduate school without close friends. There could be none closer than Craig Haimson and Thomas Kang. I'm honored to have shared the long ride with yinz guys at 5934 Phillips Ave. Many thanks are also due to DJ Bolger, John Connelly, John Klein, and Suzy Scherf for pulling me out of the house/lab periodically for a breath of fresh air and beer.

Finally, special thanks to Carol, Louie, and Madison for their generous patience, loving support, and understanding during the months of long hours I spent working on this - you kept me sane, happy, and balanced.

CONTENTS

Dedication	ii
Acknowledgements	iii
Abstract	xix
1 Introduction	1
1.1 Neural Adaptation and Synaptic Depression	2
1.2 Neural Activity Decreases in Behaving Humans and Non-human Primates	5
1.3 Repetition Suppression and Behavioral Priming	8
1.4 Previous Accounts of Repetition Priming	9
1.4.1 Spreading Activation Accounts	10
1.4.2 Episodic Retrieval Accounts	13
1.4.3 Distributed Connectionist Accounts	15
2 Neural Synchrony, Behavioral Priming and Processing Efficiency	27
2.1 The Importance of Metabolic Efficiency in Neural Processing	31
2.2 Mechanisms Underlying Spike Synchrony in Artificial Neural Networks	32
2.2.1 Synchrony in Networks with Excitatory or Inhibitory Synapses	36
2.2.2 Synchrony in Networks with Both Excitation and Inhibition	39
2.2.3 The Impact of Short-term Plasticity on Synchronization	42
2.3 Empirical Evidence for Neural Synchrony	45
2.4 Basic Hypothesis and Overview of Simulations: Short-term Plasticity Enhances Neural Synchrony and Processing Efficiency	51
2.4.1 Short-term Plasticity, Synchrony, Priming, and Processing Efficiency (<i>Simulations 1-5</i>)	53
2.4.2 Interaction of Short-term Plasticity and Neuromodulation (<i>Simulations 6-8</i>)	54

3	Spiking Neuron Simulations of the Relationships Among Repetition Suppression, Neural Synchronization, and Behavioral Priming	57
3.1	General Simulation Methods: <i>Simulations 1-5</i>	58
3.1.1	Synaptic Currents	61
3.1.2	Firing-rate Adaptation and Synaptic Depression	65
3.1.3	Stimuli.	69
3.1.4	Measuring Firing Rate and Synchrony	71
3.1.4.1	Firing Rate	71
3.1.4.2	Synchrony	71
3.2	Simulation 1: The Effect of Stimulus Repetition on Firing Rate	77
3.2.1	Short-term Repetition Suppression in Anesthetized Monkeys (Miller, Gochin, & Gross, 1991)	77
3.2.2	Short-term Repetition Suppression in Awake Monkeys Performing a Delay Match-to-Sample (DMS) task (Miller, Li, & Desimone, 1993)	81
3.2.3	Short-term Repetition Suppression in Human Adults Performing Delay Match-to-Sample (Jiang, Haxby, Martin, Ungerleider, & Parasuraman, 2000)	93
3.2.4	Short-term Repetition Suppression in Human Adults During an fMR-Adaptation Paradigm (Grill-Spector & Malach, 2001)	96
3.2.5	Comment	103
3.3	Simulation 2: The Effect of Stimulus Repetition on Synchronization	105
3.3.1	The Effect of Stimulus Repetition on Spike Synchronization and its Tolerance to Input Heterogeneity	106
3.3.2	The Impact of Synaptic Delays and Interactions with Other Cortical Regions on Repetition-Related Changes in Spike Synchronization	113
3.3.3	Impact of Inter-Stimulus Interval on Repetition-Related Changes in Spike Synchronization	119
3.3.4	Comment	122
3.4	Simulation 3: Contribution of Inhibition, Adaptation, and Synaptic Depression to Synchronization	126
3.4.1	Effect of Blocking Inhibition, Adaptation, and Synaptic Depression on Firing Rate and Synchrony	126
3.4.2	Contributions of Inhibition, Adaptation, and Synaptic Depression to Synchrony that are Unrelated to Changes in Firing Rate	129
3.4.3	Comment	132
3.5	Simulation 4: Effect of Firing Rate and Synchronization on Reaction Time.	133
3.5.1	Effect of Input Firing Rate, Input Synchrony, and Balanced Excitation/Inhibition on Output Firing Rate	138
3.5.2	Effect of Input Firing Rate and Synchrony on Reaction Time	144
3.5.3	Estimating Reaction Times for Simulation 2 (Section 3.3)	147

3.5.4	Comment	148
3.6	Simulation 5: Quantification of Efficiency	150
3.6.1	Comment	153
4	Firing-Rate and Spiking Neuron Simulations of the Interaction Between Short-Term Plasticity and Neuromodulation	155
4.1	Semantic Knowledge and the Distinction Between Access/Refractory and Degraded-Store Semantic Impairments	159
4.1.1	Access/Refractory and Degraded-Store Semantic Impairments: Patient Data	162
4.2	General Simulation Methods: Simulations 6-7	172
4.2.1	Network Architecture	174
4.2.2	Synaptic Depression	175
4.2.3	Neuromodulation	179
4.2.4	Training Procedure	183
4.2.4.1	Input Patterns	183
4.2.4.2	Target Patterns	183
4.2.4.3	Time Course of a Single Training Pattern	184
4.2.5	Lesioning Procedure	185
4.2.6	Testing Procedure	185
4.3	Simulation 6: Accounting for the Basic Contrast Between Access/Refractory and Degraded-Store Patient Performance	186
4.3.1	Rate Effects	194
4.3.2	Semantic Relatedness Effects	194
4.3.3	Frequency Effects	195
4.3.4	Serial Position Effects	196
4.3.5	Consistency Effects	197
4.4	Simulation 7: Accounting for Exceptions to the Access/Refractory and Degraded-Store Performance Patterns	201
4.4.1	Frequency Effects without Consistency	202
4.4.2	Consistency without Frequency Effects	202
4.4.3	Access/Refractory Pattern without Serial Position Effect	205
4.4.4	Comment: Simulations 6 and 7	208
4.5	Simulation 8: Dependence of Priming vs. Habituation Effects on Neuromodulatory Level in Spiking Neural Networks	216
4.5.1	Effect of Neuromodulation on Firing Rate and Synchrony	219
4.5.2	Effect of Neuromodulation on Reaction Time and Efficiency	222
4.5.3	Comment	226
5	Conclusions and Future Directions	228
5.1	Conclusions	228
5.2	Future Directions	239

References	246
Appendix A	275
Appendix B	296
Appendix C	299
Appendix D	302

LIST OF TABLES

4.1	Summary of "degraded-store" (above line) and "access/refractory" (below line) patient data.	166
4.2	Effect of serial position in the model and in patient data from Warrington & Cipolotti (1996).	197
4.3	Effect of response consistency in the model and in patient data from Warrington & Cipolotti (1996).	198

LIST OF FIGURES

1.1	<p>Evidence that neural firing rates (spikes/sec) return to baseline just after the offset of a stimulus: (A) Extracellular recordings of a monkey inferotemporal cortex neuron responding to 9 repeated presentations of a visual object (separated by delays of 20 seconds), as reported by Miller, Gochin, and Gross (1991). Firing rates (in both rasters and PSTH) return to baseline levels after the offset of the stimulus (duration of 1 sec). 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, <i>Visual Neuroscience</i>, 7, p. 359. Copyright 1991 by Cambridge University Press. Permission pending. (B) Firing rate of a complex cell in monkey V1 to repeated presentations of a stationary sinusoidal grating (averaged over 40 runs), as reported by Muller, Metha, Krauskopf, and Lennie (1999). Firing rates return to baseline rapidly after the offset of each of the 3 stimuli shown. From "Rapid adaptation in visual cortex to the structure of images," by J. Muller, A. Metha, J. Krauskopf, and P. Lennie, 1999, <i>Science</i>, 285, p. 1405. Copyright 1999 by the American Association for the Advancement of Science. Permission pending.</p>	12
1.2	<p>(A) A combination of unit activity increases and decreases induced by gradual learning lead the dot product of target and actual unit activities to increase with practice (compare familiar with unfamiliar patterns), as reported by McClelland and Rumelhart (1985, p. 175). Processing is shown as a function of time during an individual stimulus presentation. (B) Throughout learning, the number of processing cycles required to reach a fixed criterion of accuracy decreases for both familiar and unfamiliar stimuli (McClelland and Rumelhart, 1985, p. 179). From "Distributed memory and the representation of general and specific information," by J. McClelland and D. Rumelhart, 1985, <i>Journal of Experimental Psychology: General</i>, 114. Copyright 1985 by the American Psychological Association, Inc. Permission pending.</p>	18

1.3	<p>(A) Firing rates of neurons in prefrontal cortex of monkeys viewing familiar or novel visual objects, as reported by Rainer and Miller (2000). Firing rates were lower throughout stimulus processing for familiar relative to novel stimuli. From "Effects of visual experience on the representation of objects in the prefrontal cortex," by G. Rainer and E. Miller, 2000, <i>Neuron</i>, 27, p. 181. Copyright 2000 by Cell Press. Permission pending. (B) Firing rates of inferotemporal neurons in monkeys performing a delay match-to-sample task show similar decreases as a function of repetition at short delays (several seconds: Sample vs. Match) and longer delays (several minutes: 1st vs. 2nd presentation), as reported by Li, Miller, and Desimone (1993). From "The representation of stimulus familiarity in anterior inferior temporal cortex," by L. Li, E. Miller, and R. Desimone, 1993, <i>Journal of Neurophysiology</i>, 69, p.1925. Copyright 1993 by the American Physiological Society. Permission pending.</p>	21
1.4	<p>Firing rates of neurons in monkey motor cortex (FEF) during the performance of a saccade initiation task, as reported by Schall (2001). Reaction times were reliably associated with the amount of time required for movement-related cells in FEF to reach a fixed firing-rate threshold. From "Neural basis of deciding, choosing, and acting," by J. Schall, 2001, <i>Nature Reviews Neuroscience</i>, 2, p. 38. Copyright 2001 by Macmillan Magazines Ltd. Permission pending.</p>	23
1.5	<p>Short-term repetition priming effects in human subjects performing a lexical decision task, as reported by McKone (1998). Priming effects show large reductions in magnitude over several seconds, regardless of whether intervening trials are present during the prime-probe delay. From "The decay of short-term implicit memory: unpacking lag," by E. McKone, 1998, <i>Memory & Cognition</i>, 26, p.1181. Copyright 1998 by Psychonomic Society, Inc. Data re-graphed, permission pending.</p>	26
2.1	<p>Spike synchrony across a population of randomly spiking cells leads to fluctuations in the population-averaged firing rate. In contrast, asynchronous or independent random firing across the same population leads to a much more stable estimate of firing rate. The results shown are for 1000 spiking cells, each firing at 50 Hz with correlated random (partially synchronized: every cell shares identical spike times with 5% of the other cells) or independent random spike times generated from a homogenous Poisson distribution. Firing rate across the population was calculated in 10-ms bins at increments of 10 ms. The asynchronous case better meets the stationary requirement of input spiking for the derivation of firing-rate from spiking networks (e.g. Amit & Tsodyks, 1991; Gerstner, 1995; Wilson & Cowan, 1972). Firing rate networks will therefore be better approximations of spiking networks in this case.</p>	29

2.2	(A) Synchrony is unstable in Type I model cells with excitatory coupling, and (B) synchrony is stable, as is anti-synchrony in this case, for Type I cells with inhibitory coupling. The stability of particular phase-locked solutions is determined by the slope of H_{odd} where it crosses $\omega_1 - \omega_2$: Positive slope indicates stability and negative slope indicates instability (see text for explanation).	35
2.3	Heterogeneity of firing rate (indicated by non-zero $\omega_1 - \omega_2$) degrades the stability of the synchronous solution in two Type I neurons with inhibitory coupling. Since H_{odd} is always 0 at the beginning and end of the period of firing, it cannot be equal to $\omega_1 - \omega_2$. It should be noted that this depiction is only for illustrative purposes and is not precise because the calculation of H_{odd} here actually assumes equal firing rates.	38
2.4	Network architecture assumed in analyzing the interaction of coupled excitatory (e) and inhibitory (i) neuron pairs (e.g. Ermentrout et al., 2001). . .	40
2.5	Short-term repetition of odor puffs to the antennae of a locust leads to decreased spike rates in antennal lobe neurons while at the same time leading to enhanced synchronization, as reported by Stopfer and Laurent (1999): (A) local field potentials (LFPs) and voltage traces of projection neurons (PNs) initially exhibit relatively random timing relationships, yet after 12-15 odor puffs, spikes in the PNs are locked to the peaks of the LFP; (B) Firing rates in the PNs decrease across repeated odor puffs, while simultaneously showing greater coherence with the LFP - indicating better spike synchrony across the population. From "Short-term memory in olfactory network dynamics," by M. Stopfer and G. Laurent, 1999, <i>Nature</i> , 402, p. 664. Copyright 1999 by Macmillan Magazines Ltd. Permission pending.	49
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.	59
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.	61
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.	66
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.	68

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.	70
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, <i>Journal of Computational Neuroscience</i> , 5, p. 7. Copyright 1998 by Kluwer Academic Publishers. Permission pending.	73
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).	76
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, <i>Visual Neuroscience</i> , 7, p. 360. Copyright 1991 by Cambridge University Press. Permission pending.	79
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, <i>Journal of Neuroscience</i> , 13, p. 1461. Copyright 1993 by the Society for Neuroscience. Permission pending.	81

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, <i>Journal of Neuroscience</i> , 13, p. 1465. Copyright 1993 by the Society for Neuroscience. Permission pending.	84
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, <i>Journal of Neuroscience</i> , 13, p. 1467. Copyright 1993 by the Society for Neuroscience. Permission pending.	85
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, <i>Journal of Neuroscience</i> , 13, p. 1464. Copyright 1993 by the Society for Neuroscience. Permission pending.	87
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 negative 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.	89

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, <i>Journal of Neuroscience</i> , 13, p. 1466. Copyright 1993 by the Society for Neuroscience. Permission pending.	90
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, <i>Journal of Neuroscience</i> , 13, p. 1463. Copyright 1993 by the Society for Neuroscience. Permission pending.	92
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, <i>Science</i> , 287, p. 644. Copyright 2000 by the American Association for the Advancement of Science. Permission pending.	94
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, <i>Acta Psychologica</i> , 107, p. 298. Copyright 2001 by Elsevier Science B.V. Permission pending.	98

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, <i>Acta Psychologica</i> , 107, p. 303. Copyright 2001 by Elsevier Science B.V. Permission pending.	100
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 in as a function of the number of repetitions used in each condition rather than as a function of number of stimuli; (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, <i>Acta Psychologica</i> , 107, p. 305. Copyright 2001 by Elsevier Science B.V. Permission pending.	102
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.	108
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.	110
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 (1-250 excitatory and 251-300 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).	112
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.	114

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.	118
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, <i>Memory & Cognition</i> , 26, p.1181. Copyright 1998 by Psychonomic Society, Inc. Data re-graphed, permission pending.	121
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.	128
3.27	The effect of blocking inhibition (I_{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.	131
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.	141
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: $g_{EI}=0.15$), whereas (B) shows the same results for 20% adaptation and synaptic depression. See text for explanation.	143

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.	145
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.	149
3.32	Changes in a simple measure of metabolic inefficiency, the average number of spikes in each cell prior to a response, are shown for the four levels of input heterogeneity explored in <i>Simulations 2-4</i> (+/- 0, 20, 40, and 60%). Inefficiency was reduced similarly for each heterogeneity condition across stimulus repetitions.	151
3.33	Changes in the scale-free measure of metabolic efficiency, Δ Efficiency, across stimulus repetitions are shown for the four levels of input heterogeneity explored in <i>Simulations 2-4</i> (+/- 0, 20, 40, and 60%). Efficiency was increased similarly for each heterogeneity condition; changes in all cases are relative to the firing rates and reaction times to Stimulus 1. . .	153
4.1	Network architecture.	174
4.2	Effect of time-varying pre-synaptic firing rate on synaptic depression in the spiking Varela et al. (1999) model and the firing-rate connectionist approximation.	178
4.3	Pre-synaptic effect of neuromodulation on transmitter release and the buildup of synaptic depression. Depicted on the y-axis is the post-synaptic net input (η) for a single synapse (equal to pre-synaptic activity, sinusoidally modulated, multiplied by a weight of 1.0 and scaled by the release factor, $\rho(M)$, and the depression factor, D). Two different levels of neuromodulation are shown for comparison, very low ($M=-4.0$) and moderately high ($M=2.0$).	180
4.4	Effect of neuromodulation on post-synaptic activity. Shown on the y-axis is unit activity, $a(\eta)$ (function of net input, η , with the gain, $g(\)$, dependent on neuromodulation, M). The increase in gain/sensitivity is apparent for moderately high levels of neuromodulation ($M=2.0$) compared to very low levels ($M=-4.0$).	182
4.5	Effect of rate in the model under different damage combinations and the same effect in the patient data. The values of M and % lesioned connections are listed for each damage combination.	188
4.6	Effect of semantic relatedness in the model under different damage combinations and the same effect in the patient data.	189
4.7	Effect of frequency in the model under different damage combinations and the same effect in the patient data.	191

4.8	Effect of serial position in the model under different damage combinations (fast presentation rate conditions, RSI = 1 sec).	192
4.9	Surface plots showing the directions and magnitudes of the <i>rate</i> , <i>semantic relatedness</i> , <i>frequency</i> , and <i>serial position</i> effects for the entire space of damage combinations. For each effect, the x- and y-axes represent the two damage types (neuromodulation, <i>M</i> , and damage to connections, % <i>Lesion</i>), and the z-axis represents the magnitude of the effect in terms of proportion correct (e.g. the slow minus the fast condition for the rate effect, distant - close for the semantic relatedness effect, etc.). The magnitudes of the effects are also represented in grayscale, where white corresponds to the maximum effect value and black corresponds to the minimum value; colorbars to the right of each plot indicate scale.	207
4.10	The impact of stimulus repetition and neuromodulatory level, <i>M</i> , on firing rate means (A) and firing rate standard deviations (B). Stimulus repetition reduced both the means and standard deviations of firing rates, although rates were larger overall for higher levels of <i>M</i>	220
4.11	Impact of stimulus repetition and neuromodulatory level, <i>M</i> , on coherence values among the excitatory cells. Coherence increased strongly across repetitions for moderate values of <i>M</i> but not for low values (little or no increases for $M < -1.0$). For larger values of <i>M</i> , coherence increases were not as robust as for moderate values (e.g. $M = +2.0$).	222
4.12	The impact of stimulus repetition and neuromodulatory level on reaction times. For low levels of neuromodulation ($M < -1.0$), stimulus repetition led to slower reaction times and behavioral habituation effects. In contrast, stimulus repetition led to faster reaction times and repetition priming for more moderate levels of neuromodulation (e.g. $M = -0.5, 0.0, \text{ and } 0.5$).	224
4.13	Changes in metabolic efficiency as a function of stimulus repetition and neuromodulatory level. Efficiency showed significant increases across repetitions for all levels of <i>M</i> , although the changes were larger for more moderate levels.	225

Abstract

Neural cells continually and automatically adjust their responsiveness to input and their impact on other cells. While these dynamic changes in sensitivity are well documented and have been broadly observed in a variety of brain regions, the functional role that these changes play in the neural circuits that underlie human perception and cognition is far from clear. In this thesis, I will argue that short-term neural plasticity mechanisms such as firing-rate adaptation and synaptic depression play a central role in affording accurate, yet metabolically efficient information processing. In the neurologically intact brain, adaptation and synaptic depression reduce neural firing rates under repeated or continued exposure to a stimulus by automatically and locally attenuating synaptic efficacy while simultaneously synchronizing neural activity - permitting coordinated, efficient signal propagation and enhanced perception. However, this benefit is limited: If firing rate reductions are too extreme, signals can attenuate to a point where perceptual abilities become impaired. Global neuromodulatory systems that are sensitive to behavioral state and performance level in a task can dynamically adjust the gain of cortical activity when performance is poor, ensuring accurate behavioral performance at the lowest metabolic cost. Under damage to these neuromodulatory systems, cortical gain is diminished and firing rate reductions can build up to the point where sensory signals are no longer reliably propagated, leading to refractory effects in behavioral performance. In a series of spiking- and firing-rate neural network simulations that instantiate basic properties of short-term plasticity and neuromodulation, I will demonstrate that the repeated processing of a stimulus can facilitate subsequent identification ("repetition priming") or impair it ("habituation"), depending on the level of neuromodulation. As long as neuromodulatory levels are not too low, identification performance can improve while neurons reduce their firing rates and synchronize more with each repetition, dynamically enhancing processing efficiency. The simulations address observations from neural recording and neuroimaging studies associated with short-term stimulus repetition, as well as common behavioral effects in neurologically intact and disordered populations.

Chapter 1

Introduction

One of the most common properties of neural cells is that they respond transiently to constant or repeated stimulation (e.g., Adrian, 1926; McCormick et al., 1985). For example, excitatory neurons in the cortex of the mammalian brain - neurons thought to play a central role in the biological implementation of perceptual and cognitive processing - temporarily reduce or "adapt" their sensitivity to input following depolarizing pulses and spiking activity (e.g., Madison & Nicoll, 1984; Schwindt et al., 1988). Similarly, cortical cells reduce their impact on other cells following spiking activity by decreasing temporarily the amount of transmitter that they release, a phenomenon referred to as short-term "synaptic depression" (e.g., Abbott et al., 1997; Tsodyks & Markram, 1997). While these dynamic changes in sensitivity appear to occur across a wide variety of different brain regions, their functional role in perceptual and cognitive processing is unclear. In this thesis, I will describe a series of computational simulations that attempt to relate these transient changes in neural sensitivity to transient

changes in human perceptual abilities. The view that emerges from the work is that the mechanisms underlying decreases in neural responses serve to minimize the metabolic cost associated with neural information processing. When combined with neuromodulatory mechanisms that are capable of counteracting response decreases when necessary, efficiency and accuracy can be optimized simultaneously.

1.1 Neural Adaptation and Synaptic Depression

When excitatory cortical cells are stimulated to fire spikes with a constant applied current, they fire vigorously initially and then fire progressively slower. This effect of firing-rate *adaptation* has been studied extensively, and much about the underlying cellular mechanisms has been uncovered (e.g., Ahmed, Allison, Douglas, Martin, & Whitteridge, 1998; Avoli & Olivier, 1989; Connors et al., 1982; Constanti & Sim, 1987; Lorenzon & Foehring, 1992; Madison & Nicoll, 1984; McCormick et al., 1985; Sah, 1996; Sanchez-Vives et al., 2000; Schwindt et al., 1988, 1989, 1992). In most cases, decreases in excitability have been mediated by the activation of potassium (K⁺) currents that hyperpolarize the membrane potential and increase the conductance of the membrane, reducing the impact of synaptic inputs and yielding lower spiking rates. Some of these currents activate and decay quite rapidly such as the M-current, a voltage-dependent K⁺ current that opens at the higher membrane potentials that are reached during spike events, activating over a few milliseconds and decaying over tens of milliseconds (e.g. Adams, Brown, & Constanti, 1982; Constanti & Galvan, 1983). Other K⁺ currents are activated instead by the influx of ions during spiking and can be considerably slower, lasting hundreds of milliseconds in the case of calcium-activated K⁺

currents ($I_{K+(Ca^{2+})}$: Constanti & Sim, 1987; Liu & Wang, 2001; Madison & Nicoll, 1984) or even tens of seconds in the case of K^+ currents activated by Na^+ ($I_{K+(Na^+)}$: Schwindt et al., 1989; Vives-Sanchez et al., 2000). Each of these currents is observed in excitatory cells throughout the cortex (e.g. pyramidal cells), but are notably absent in inhibitory interneurons (Azouz et al., 1997; McCormick et al., 1985). Accordingly, most inhibitory interneurons do not show firing-rate adaptation effects when stimulated with constant current¹. Furthermore, neuromodulatory agents that block these K^+ currents selectively such as acetylcholine and norepinephrine also block firing-rate adaptation effects, resulting in higher and sustained activity (e.g. Hasselmo & Bower, 1992; Lorenzon & Foehring, 1992; Madison & Nicoll, 1984; McCormick & Prince, 1986; McCormick & Williamson, 1989; Schwindt et al., 1988). These results are bolstered by biophysical simulations of individual cells showing that the inclusion of spike-dependent potassium currents is sufficient to explain observed empirical characteristics of firing-rate adaptation (e.g. Barkai & Hasselmo, 1994; Ermentrout, 1998a; Liu & Wang, 2001; Migliore et al., 1995).

Synaptic depression, the temporary reduction in synaptic efficacy following pre-synaptic activity, has also been studied in some detail, although the underlying cellular mechanisms are somewhat less clear than for firing-rate adaptation (e.g., Abbott et al., 1997; Deisz & Prince, 1989; Finlayson & Cynader, 1995; Markram & Tsodyks, 1996; Thomson et al., 1993, 1996; Tsodyks & Markram, 1997; Varela et al., 1997, 1999). Synaptic depression occurs at both excitatory and inhibitory synapses throughout the

¹ Even though most inhibitory interneurons do not show firing-rate adaptation effects when stimulated directly with current, they are still expected to show decreases in firing rate in more natural circumstances because connections between different cortical regions are nearly exclusively excitatory (White, 1989), and inhibitory cells receive their drive from excitatory cells that do exhibit firing-rate adaptation.

cortex, with larger magnitudes of depression at excitatory synapses (Galarreta & Hestrin, 1998; Varela et al., 1999). The time course of synaptic depression ranges from build-up and recovery times of several hundred milliseconds to tens of seconds at excitatory synapses and several seconds at inhibitory synapses (e.g. Finlayson & Cynader, 1995; Galarreta & Hestrin, 1998; Tsodyks & Markram, 1997; Varela et al., 1997, 1999). It appears to depend almost entirely on pre-synaptic activity rather than on post-synaptic activity, and it is mediated by reductions in pre-synaptic transmitter release (Finlayson & Cynader, 1995; Nelson, 1991; Thomson et al., 1993; Tsodyks & Markram, 1997). The exact mechanisms that give rise to the reduction in transmitter release are not known, but several possibilities have been suggested such as transmitter depletion, inactivation of the calcium current that triggers transmitter release, and the down-regulation of release by pre-synaptic autoreceptors (e.g. McLean & Palmer, 1996; Senn, Markram, & Tsodyks, 2001; Varela et al. 1999). There is some evidence that synaptic depression actually depends on transmitter release itself, rather than on pre-synaptic spiking (Tsodyks & Markram, 1997). Synapses for which transmitter release is less likely (given pre-synaptic spiking) undergo less depression. Consistent with this, neuromodulators like acetylcholine and norepinephrine that reduce transmitter release at cortical synapses also reduce the magnitude of synaptic depression effects (e.g. Tsodyks & Markram, 1997; Gil, Connors, & Amitai, 1997). Taken together with the effects of neuromodulation on firing-rate adaptation, these findings suggest that the transient firing rate decreases due to synaptic depression and adaptation may be diminished under behavioral circumstances that lead to heightened release of acetylcholine and/or norepinephrine.

1.2 Neural Activity Decreases in Behaving Humans and Non-human Primates

In behavioral circumstances that involve the repeated identification of stimuli over a short period of time, decreases in neural activity are commonly observed. For example, Jiang et al. (2000) demonstrated in a recent event-related fMRI study with human participants that during performance of a delayed-match-to-sample task, blood flow (and presumably neural activity) in ventral temporal, occipital and parietal cortices was reduced within a block of stimuli for repeated items. This blood flow decrease was stimulus specific, built up over 4 to 5 stimulus repetitions, occurred regardless of whether the repeated stimuli were targets or distractors, and largely recovered between blocks of trials. Similar response decrements have been observed following stimulus repetition in both fMRI (e.g. left inferior prefrontal cortex: Demb et al., 1995; Wagner, Maril, & Schacter, 2000; primary visual, inferotemporal, dorsal extrastriate, and posterior parietal cortices: Buchel, Coull, & Friston, 1999; Grill-Spector & Malach, 2001) and PET (e.g. posterior ventral occipitotemporal and parietal cortices: Buckner et al., 1995; Schacter, Alpert, Savage, Rauch, & Albert, 1996; Squire et al., 1992). In some cases, a significant response decrement has been shown to last as long as three days, although this long-lasting decrement is somewhat smaller than that observed following immediate repetition (e.g. van Turennout, Ellmore, & Martin, 2000; Wagner et al., 2000). The close association of stimulus repetition with reduced neural activity has led some researchers to refer to this phenomenon as *repetition suppression* (Desimone, 1996; Wiggs & Martin, 1998). This is not to say that stimulus repetition is always associated with decreased neural activity. In some instances, activity increases have been observed, notably in medial temporal and

anterior prefrontal cortices that are most associated with explicit and intentional recollection of prior events and episodes (e.g. Schacter & Buckner, 1998; Schacter, Buckner, & Koutstaal, 1998).

The phenomenon of repetition suppression was, in fact, first described in extracellular single-neuron recording studies of non-human primates (e.g. Baylis & Rolls, 1987; Li, Miller, & Desimone, 1993; Miller, Li, & Desimone, 1991; 1993; see Brown & Xiang, 1998; Desimone, 1996, for reviews). Similarly to effects in blood flow documented in neuroimaging studies, the firing rate of individual neurons shows a marked decrease following stimulus repetition. This decrease is stimulus-specific (e.g. Miller et al., 1991, 1993), is greater for stimuli that initially evoke higher firing rates (Li et al., 1993), accumulates over 5-7 stimulus repetitions (Miller, Gochin, & Gross, 1991), and largely recovers following a short delay, although longer-lasting effects have also been observed (e.g. Miller et al., 1993; Li et al., 1993; Rainer & Miller, 2000). Repetition suppression appears to be very automatic in that it can occur in a delay-match-to-sample task, during passive fixation, under anesthesia, and it survives a cholinergic blockade (e.g. Miller & Desimone, 1993; Miller et al., 1991). While it has been studied primarily in inferotemporal (IT) cortex, similar effects have been observed in a variety of other cortical regions (e.g. prefrontal cortex: Miller, Erikson, & Desimone, 1996; primary visual cortex: Muller, Metha, Krauskopf, & Lennie, 1999; extrastriate cortex: Haenny & Schiller, 1988; perirhinal cortex: Brown & Xiang, 1998). Furthermore, such effects do not appear to be explained solely by peripheral sensory adaptation (Lueschow, Miller, & Desimone, 1994).

Interestingly, the particular characteristics of short-term repetition suppression in neuroimaging and single-neuron recording studies are broadly consistent with the properties of firing-rate adaptation and synaptic depression discussed above. The stimulus-specific nature of repetition suppression can be explained by the cell- and synapse-specificity of firing-rate adaptation and synaptic depression: Only those cells and synapses that have been recently activated in the processing of a previous stimulus will undergo adaptation and synaptic depression. Cells and synapses involved in the processing of other stimuli will be unaffected. Several studies of the time course of firing-rate adaptation and synaptic depression have measured recovery time constants of 10 seconds or longer for excitatory cells/synapses and 2 to 4 seconds for inhibitory synapses, suggesting that complete recovery can require as much as a minute or more (e.g. Finlayson & Cynader, 1995; Galarreta & Hestrin, 1998; Sanchez-Vives et al., 2000; Schwindt et al., 1989; Varela et al., 1997; 1999). If stimuli are repeated within the time window of recovery, adaptation and synaptic depression can build up with repetition. Like repetition suppression effects, adaptation and synaptic depression appear to be very automatic in that they require only spiking activity to occur, and they are both commonly observed in experiments with cortical slices that have been removed from the brain. The degree of adaptation and synaptic depression observed is roughly proportional to firing rate (i.e. the greater the firing rate, the stronger the adaptation and depression; for detailed model fits, see Abbott et al., 1997; Barkai & Hasselmo, 1994; Tsodyks & Markram, 1997; Varela et al., 1997, 1999; Wang, 1999). One would also expect firing-rate adaptation and synaptic depression to be present under a cholinergic blockade as observed by Miller and Desimone (1993), because any cholinergic suppression of

adaptation and synaptic depression would be prevented. In contrast, the longer-lasting effects of repetition suppression on the orders of several minutes, hours or days could not be explained by adaptation and synaptic depression. These effects are presumably underlied by longer-lasting plasticity changes such as long-term potentiation (LTP) and long-term depression (LTD), mechanisms that depend jointly on pre- and post-synaptic activity states (e.g. Bi & Poo, 1998; Bliss & Gardner-Medwin, 1973; Bliss & Lomo, 1973; Kirkwood, Rioult, & Bear, 1996; Markram, Lubke, Frotscher, & Sakmann, 1997; see Malenka & Nicoll, 1999, and Rioult-Pedotti, Friedman, & Donoghue, 2000, for recent discussions).

1.3 Repetition Suppression and Behavioral Priming

While neural activity commonly decreases with repeated stimulus processing, behavioral performance commonly improves. People are often faster and more accurate at identifying a stimulus if it has been recently processed, a phenomenon referred to as *repetition priming* (e.g. Cave & Squire, 1992; Ratcliff & McKoon, 1988; Scarborough et al., 1977). Repetition priming effects can be very long-lasting, surviving delays of up to 48 weeks (e.g. Cave, 1997), although there are large reductions over short delays of seconds (e.g. McKone, 1995, 1998), and from seconds to days (e.g. van Turennout et al., 2000). Priming can be incremental across a number of repetitions, yielding slight decreases in reaction time with each repetition (e.g. Brown et al., 1996; Logan, 1990; Ostergaard, 1998; Wiggs et al., 1997). These effects are very automatic in the sense that they can occur when subjects are anesthetized during the initial stimulus presentation

(Jelicic, Bonke, Wolters, & Phaf, 1992; Kihlstrom & Schacter, 1990), they do not require explicit awareness that the stimulus was presented before (e.g. Cave & Squire, 1992; Tulving et al., 1982), and they are robust to attentional manipulations (e.g. Kellogg et al., 1996; Szymanski & MacLeod, 1996) and small alterations of stimulus form (e.g. Biederman & Cooper, 1991; 1992; Cave et al., 1996; Srinivas, 1996). These properties of priming appear broadly consistent with the characteristics of repetition suppression, and indeed, the magnitude of repetition priming effects in neuroimaging studies has been shown to be correlated with the magnitude of repetition suppression effects (e.g., van Turennout et al., 2000; Wagner, Maril, & Schacter, 2000). The close relationship that has been observed between stimulus repetition, repetition priming and repetition suppression has led some researchers to propose that the mechanisms underlying reduced neural responses also underlie the observed behavioral facilitation (e.g. Desimone, 1996; Wiggs & Martin, 1998). In particular, Desimone (1996) has suggested that the joint observations of reduced activity and enhanced identification performance may reflect the "sharpening" of representations such that individual cells are encoding stimuli more selectively with an overall reduction in the number of active cells.

1.4 Previous Accounts of Repetition Priming

While repetition suppression effects appear to be correlated with repetition priming effects, it is unclear how firing rate decreases can lead to faster, more accurate perceptual and cognitive processing. The neural observations are in some sense counterintuitive in that cognitive-level explanations of priming have generally involved mechanisms that lead to sustained or enhanced activity. It is important not to underestimate the gap that

currently exists between the notion of activity in cognitive-level models and activity in neural/biophysical models, but some basic similarities exist that permit a cursory evaluation of the degree to which mechanisms developed in cognitive psychology appear to be expressed in the function of real neural circuits (e.g. larger inputs lead to higher activity; higher activity results in larger impact on subsequent processing; short-term residual activity can occur; longer-term changes in processing can be expressed by alterations to links/connections/synapses; etc.). Most previous cognitive-level accounts of priming fall into three broad classes that will be discussed in turn: 1) spreading activation, 2) memory retrieval, and 3) distributed connectionist.

1.4.1 Spreading Activation Accounts

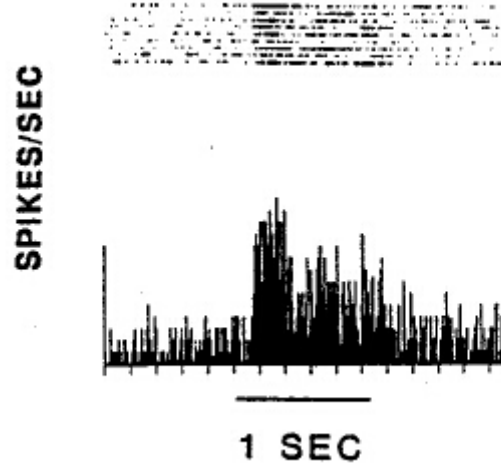
In spreading activation theories, priming results from slowly decaying residual activity or decreased activity thresholds in nodes that represent stimuli in long-term memory (e.g., Anderson, 1976, 1983; Cohen & Dehaene, 1998; Collins & Loftus, 1975; Morton, 1969, 1979; Neely, 1977; Quillian, 1967). This means that when a stimulus is processed a second time, less additional activity is required to reach threshold, and a response is initiated earlier. However, neural recording experiments involving repeated stimulus processing have not found much evidence for residual neural activity following the generation of a response unless the task itself is explicitly designed as a working memory task (e.g., Miller, Erickson, & Desimone, 1996; Rao, Rainer, & Miller, 1997). Neural responses in posterior cortical regions thought to be critically involved in perceptual processing generally last less than one second in passive fixation or perceptual tasks without explicit trial-to-trial dependencies, after which activity returns to a baseline level

(e.g., Miller, Gochin, & Gross, 1991; Muller et al., 1999; Reynolds, Chelazzi, & Desimone, 1999). Baseline neural activity occurring between individual trials also appears to be stable and unchanging throughout the course of an experimental session when stimuli are being repeated (e.g. Li et al., 1993; Miller et al., 1993). This means that the time course of residual activity in real neurons is too fast to explain short-term priming effects that span several seconds and a number of intervening stimuli in simple identification/discrimination tasks (e.g. word naming and lexical decision: McKone, 1995, 1998)². Furthermore, there is no obvious correlate of an activity response threshold that triggers the spread of activity aside from the voltage threshold for action potential generation - a threshold that is reached numerous times for each cell during the processing of any stimulus and that bears little relation to the termination of processing (each spike is automatically communicated to subsequent cells, and refractory periods last approximately 1-2 milliseconds afterward)³. Figure 1.1 illustrates several of these points.

² On the other hand, residual neural activity could well explain the short-lived priming effects in masked priming paradigms that span several hundred milliseconds (e.g. Forster & Davis, 1984; Forster, Davis, Schoknecht, & Carter, 1987; Humphreys, Besner, & Quinlan, 1988).

³ If there were a way to implement a global "reset", such that when activity in a certain population of neurons hit a critical threshold, it might be possible to account for joint reductions in activity and enhanced performance by assuming a progressive lowering of this threshold. However, in order to explain the stimulus specificity of the effects, there would need to be different thresholds for each stimulus-specific neural population, that once reached, could terminate activity (activity would have to be terminated to guarantee the activity decreases). It is not obvious what neural mechanisms could mediate such effects at the required time scales. Decreasing such a threshold would carry the added danger of higher error rates due to random activity fluctuations. The idea further predicts that neural responses should be more extended for earlier processing than for later processing, since activity takes time to propagate between regions (~ 10-15 milliseconds per region) - a prediction not obviously borne out by current data.

A.



B.

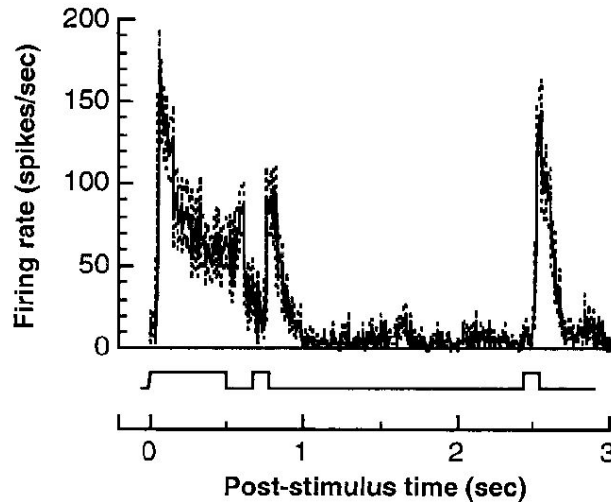


Figure 1.1: Evidence that neural firing rates (spikes/sec) return to baseline just after the offset of a stimulus: (A) Extracellular recordings of a monkey inferotemporal cortex neuron responding to 9 repeated presentations of a visual object (separated by delays of 20 seconds), as reported by Miller, Gochin, and Gross (1991). Firing rates (in both rasters and PSTH) return to baseline levels after the offset of the stimulus (duration of 1 sec). 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. 359. Copyright 1991 by Cambridge University Press. Permission pending. (B) Firing rate of a complex cell in monkey V1 to repeated presentations of a stationary sinusoidal grating (averaged over 40 runs), as reported by Muller, Metha, Krauskopf, and Lennie (1999). Firing rates return to baseline rapidly after the offset of each of the 3 stimuli shown. From "Rapid adaptation in visual cortex to the structure of images," by J. Muller, A. Metha, J. Krauskopf, and P. Lennie, 1999, *Science*, 285, p. 1405. Copyright 1999 by the American Association for the Advancement of Science. Permission pending.

1.4.2 Episodic Retrieval Accounts

The second class of priming theories involves the retrieval of individual episodes related to the processing of the prime stimulus or other prior stimuli (e.g. Doshier & Rosedale, 1989; Jacoby, 1983; Logan, 1988, 1990; McKoon & Ratcliff, 1992; Ratcliff & McKoon, 1988; Whittlesea & Jacoby, 1990). For example, in the compound-cue theory of Ratcliff and McKoon (1988), many psychophysical tasks involving a binary decision (e.g., lexical decision) are performed on the basis of familiarity (see also Balota & Chumbley, 1984). On this view, the current stimulus is combined with the episode of the previous stimulus to form a compound cue that is used to access memory. To the extent that the current stimulus is similar to recent stimuli (e.g. identical), the returned familiarity value is higher, facilitating the decision and decreasing reaction time. Another variant of the episodic retrieval idea has been proposed by Logan (1988, 1990) in his instance-based theory of automaticity. This theory holds that a horse race ensues between slower algorithmic processing and faster episodic retrieval. Novel stimuli/tasks have no corresponding episodes in memory so processing is slower. As a stimulus is repeatedly encountered, each separate instance is stored in memory, and there is a higher and higher chance that the faster memory retrieval will win the horse race, facilitating reaction times with practice. Because the critical mechanisms of these theories have not been specified directly in terms of activity dynamics, it is not clear how to relate them to neural activity. If one were to assume a direct relationship between neural activity and familiarity or the probability of episodic retrieval, decreases in neural activity with repeated practice would clearly be inconsistent with these theories, as performance should degrade rather than improve. It is interesting to note on this point that successful memory retrieval for

previous episodes as measured in "explicit" tasks (tasks that require an explicit reference back to a prior processing episode, such as cued recall, recognition memory, etc.) has been associated with higher rather than lower neural activity in anterior prefrontal and medial temporal cortical regions (for reviews, see Schacter & Buckner, 1998; Schacter, Buckner, & Koutstaal, 1998). Indeed, Wagner, Maril, and Schacter (2000) have demonstrated that the magnitudes of priming effects and the correlated neural activity decreases can be inversely related to correct performance on explicit memory tasks. When these results are combined with the well-known findings that amnesic patients with damage to medial temporal brain regions are impaired at consciously remembering recent events and episodes but show normal or near-normal priming effects (e.g. Cave & Squire, 1992; Cermak, Talbot, Chandler, & Wolbarst, 1985; Hamann & Squire, 1997), it appears that priming accounts that rely exclusively on episodic retrieval may be in trouble⁴. The behavioral and neural dissociations between priming effects and memory effects that occur in tasks requiring an explicit reference to prior events (e.g. recall and recognition tasks) have led many researchers to suggest that there are at least partially separable memory systems that underlie the different effects (e.g. Graf & Schacter, 1985; McClelland, McNaughton, & O'Reilly, 1995; Schacter, Buckner, & Koutstaal, 1998; Squire, 1994). Nevertheless, some neuroscientists have argued for a commonality between the effects - that heightened familiarity and enhanced recognition is associated with decreased neural activity - based on neural recording evidence from perirhinal cortex in awake, behaving monkeys (see Brown & Xiang, 1998, for a review). This

⁴ While explicit memory tasks are thought to index the ability to recollect consciously prior processing episodes (containing information about the time, place, and other contextual aspects of particular experiences), it remains possible that (at least some) contextual aspects of prior episodes could be encoded in regions other than medial temporal cortices such that they could influence performance on priming tasks. It also seems likely that conscious recollection contributes to normal priming effects to a certain extent.

means that in alternative-forced-choice tasks with two alternatives, decreased activity could explain faster reaction times by removing support from one of the two possible responses that compete locally through lateral inhibition (e.g. decreasing input to neurons representing the "not a word" response in lexical decision, thus reducing lateral inhibitory input to neurons representing the competing "word" response). However, if the number of response alternatives is larger than two, this "discounting" form of explanation breaks down because eliminating one alternative does not help much with deciding amongst the remaining alternatives. The fact that neural activity decreases are correlated with priming in tasks such as object naming (e.g. van Turennout et al., 2000) undermines the generality of this type of account.

1.4.3 Distributed Connectionist Accounts

Distributed connectionist theories of priming have explained faster reaction times by positing a combination of different mechanisms, including residual activity and longer-term learning (e.g. Becker, Moscovitch, Behrmann, & Joordens, 1997; Masson, 1991, 1995; McClelland & Rumelhart, 1985; McRae, de Sa, & Seidenberg, 1997; Plaut, 1995; Plaut & Booth, 2000; Sharkey, 1990; Stark & McClelland, 2000). Distributed connectionist models bear some similarities to both the spreading activation and episodic retrieval models discussed above in that activity is communicated between units and unique experiences can lead to unique learning-related changes, although they differ from these previous models in a number of important respects. Rather than individual units representing particular perceptual features or concepts with direct psychological interpretations ("localist" representations), each stimulus is represented as a distributed

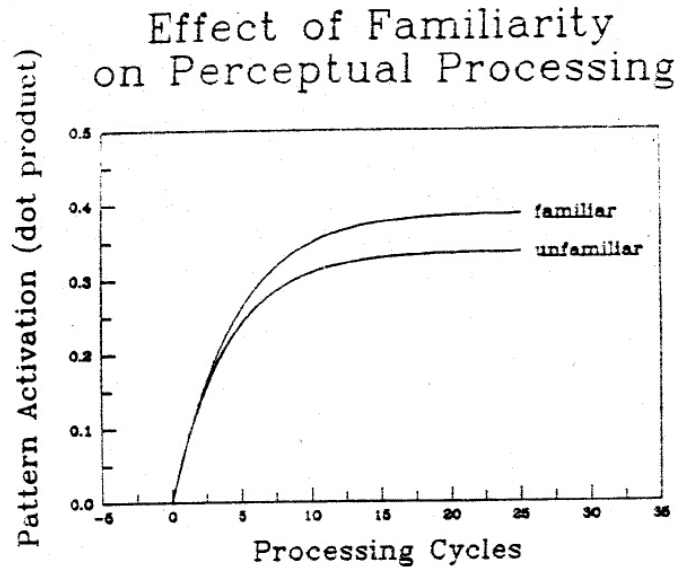
pattern of activity across all of the units in the network. Representations in these models, rather than being stipulated, develop gradually as the models learn to map specified input patterns to output patterns, and the distributed nature of the developing representations allows them to generalize knowledge automatically to untrained patterns that are similar to known patterns. This differs markedly from episodic retrieval models in which memories are retained as separate traces in an independent store: Learning and processing take place within the same system. Reaction times are often simulated as the number of computing cycles it takes for unit activities to reach some criterion measure of accuracy or to reach stable values in networks with recurrent interactions among the units.

Several connectionist models have been proposed to account for short-term priming effects by assuming that once unit activities have stabilized for a given input pattern, they stay stable until the next input pattern is presented. If two sequentially presented stimuli are similar in that they share many of the same active units, only a few units need to change their states in order to move from the first stable state into the second, yielding faster reaction times than if they shared few or no active units (e.g., Masson, 1991, 1995; McRae, de Sa, & Seidenberg, 1997; Sharkey, 1990). This idea is similar in principle to the early model of Rips, Shoben, and Smith (1973), in which similar stimuli were represented "closer" to each other in long-term memory by sharing many of the same conceptual features (e.g., has four legs, is a mammal, etc.); the representations are closer to each other in the sense that they correspond to points that are near each other in the multidimensional space defined by the various features. However, as discussed above, theories that rely completely on residual activity states in order to account for priming effects that span several seconds appear to be inconsistent with

observations that residual activity in real neurons commonly does not last this long (e.g. Figure 1.1). Aside from this point, models that settle to a stable state and stay there have difficulty addressing priming effects that span the processing of intervening stimuli because each previous state is erased by the processing of the subsequent stimulus (for discussion, see Joordens & Besner, 1992; Masson, 1995; Plaut, 1995).

Other distributed connectionist models have used longer-lasting changes to connection strengths that occur during learning in order to address longer-term priming effects that can span intervening stimuli (e.g., Becker et al., 1997; McClelland & Rumelhart, 1985; Plaut, 1995; Plaut & Booth, 2000; Stark & McClelland, 2000). The fundamental idea is that processing and learning go hand in hand. Each time a stimulus is processed, the weighted connections between units are adjusted slightly by a learning algorithm so that processing becomes more accurate. As stimuli are repeatedly encountered, the output unit activities for each stimulus gradually approach the "target" or "teacher" states specified by the training patterns (using error-correcting learning algorithms such as the delta rule with backpropagation, e.g. Rumelhart, Hinton, & Williams, 1986). In networks with continuous temporal dynamics and recurrent interactions among units, this gradual tuning of connection strengths throughout learning also leads to faster settling times, producing priming effects. This is depicted graphically in Figure 1.2.

A.



B.

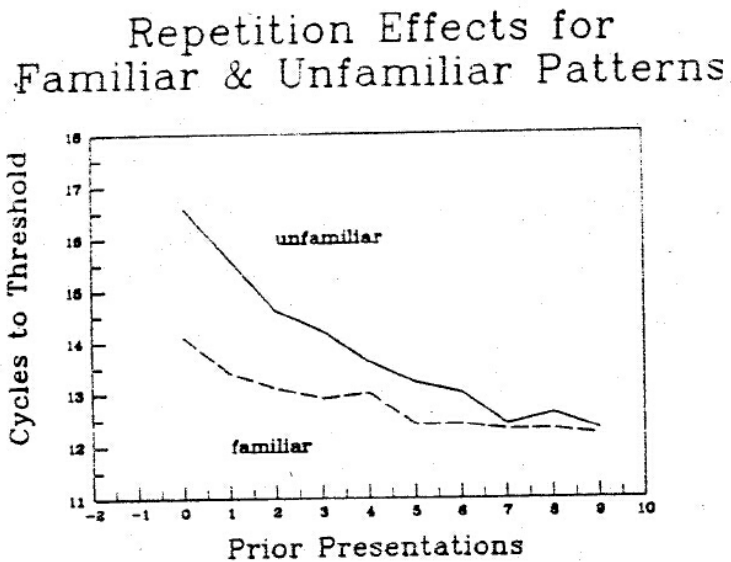


Figure 1.2: (A) A combination of unit activity increases and decreases induced by gradual learning lead the dot product of target and actual unit activities to increase with practice (compare familiar with unfamiliar patterns), as reported by McClelland and Rumelhart (1985, p. 175). Processing is shown as a function of time during an individual stimulus presentation. (B) Throughout learning, the number of processing cycles required to reach a fixed criterion of accuracy decreases for both familiar and unfamiliar stimuli (McClelland and Rumelhart, 1985, p. 179). From "Distributed memory and the representation of general and specific information," by J. McClelland and D. Rumelhart, 1985, *Journal of Experimental Psychology: General*, 114. Copyright 1985 by the American Psychological Association, Inc. Permission pending.

In terms of the activities of individual units, the effect of training is to lead some units to reach higher activity values faster, whereas other units reach lower activity values faster. In other words, stimulus repetition leads to a combination of activity increases and decreases. In networks that employ only positive activation values with an upper bound (e.g. ranging between 0 and 1), the effect of learning is to make unit activities more binary (on/off) and selectively responsive to particular stimuli, where a certain subset of the units is critical in representing a given stimulus⁵. This is similar to the idea proposed by Desimone (1996) in which learning leads to a "sharpening" of neural responses; neurons increase their selectivity by decreasing their responses to most stimuli, yet retaining responses for a small subset of stimuli.

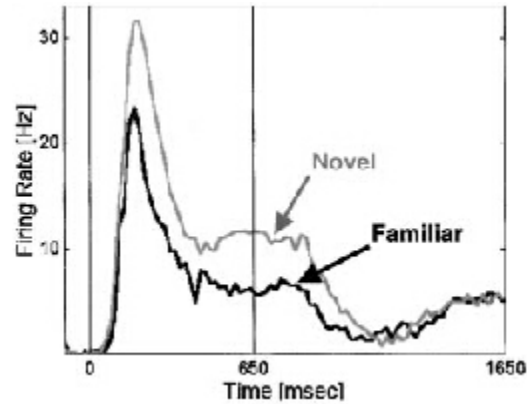
Evidence available from neuroscience studies appears to provide mixed support for the longer-term learning mechanisms used by current distributed connectionist models. Neurons in inferotemporal and prefrontal cortices of monkeys have been found to exhibit enhanced stimulus selectivity following extensive practice (e.g. Kobatake, Wang, & Tanaka, 1998; Rainer & Miller, 2000). For example, Kobatake et al. (1998) trained two macaque monkeys for several months on a task that required them to identify a sample shape stimulus that was presented several seconds earlier out of an array of 5 stimuli. Following training, the two monkeys and 3 control monkeys were anesthetized, and neural responses in area TE of inferotemporal cortex to individual trained and novel reference stimuli were recorded. Cells fired at higher rates to training stimuli in the trained monkeys than in control monkeys, while the distributions of firing rates to the novel reference stimuli were identical for trained and control monkeys. Cells in all

⁵ This follows from the fact that a unit with an activity of 0 sends no output to other units, whereas a unit with an activity of 1 will tend to send a large amount of output. Units with larger activities also tend to build larger outgoing connections in these models.

monkeys tended to be broadly tuned to multiple stimuli. The effect of training on the stimulus selectivity of the neural responses was assessed by treating the firing rates elicited by stimuli as vectors of activities across the recorded cells. The distances between pairs of activity vectors corresponding to training stimuli were found to be greater in trained than in control monkeys, indicating that the firing-rate responses to these stimuli were more discriminable and that individual cells were relatively more selective for particular training stimuli. While Rainer and Miller (2000) have similarly found enhanced stimulus selectivity in monkey prefrontal neurons following repeated practice in a delay match-to-sample task, they found little evidence of practice-related increases in firing rate - only practice-related decreases; across a variety of studies, this has been the more common pattern (see Brown & Xiang, 1998; Desimone, 1996, for review/discussion). Figure 1.3 shows examples of the detailed time course of these decreases from Rainer and Miller (2000) and from a similar study of inferotemporal cortex neurons by Li, Miller, and Desimone (1993). Note that the peak values of firing rate are lower following repetition and that the onset of the neural response does not appear to be advanced⁶.

⁶ It is also of note that these results are directly at odds with the arguments put forward by James et al. (2000) in explaining the relationship between repetition suppression and behavioral priming (to the extent that results from non-human primate studies are applicable to understanding similar results obtained in studies with humans).

A.



B.

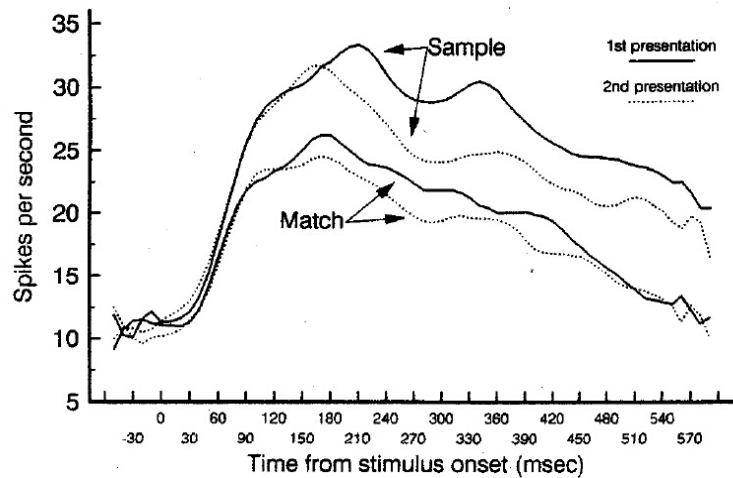


Figure 1.3: (A) Firing rates of neurons in prefrontal cortex of monkeys viewing familiar or novel visual objects, as reported by Rainer and Miller (2000). Firing rates were lower throughout stimulus processing for familiar relative to novel stimuli. From "Effects of visual experience on the representation of objects in the prefrontal cortex," by G. Rainer and E. Miller, 2000, *Neuron*, 27, p. 181. Copyright 2000 by Cell Press. Permission pending. (B) Firing rates of inferotemporal neurons in monkeys performing a delay match-to-sample task show similar decreases as a function of repetition at short delays (several seconds: Sample vs. Match) and longer delays (several minutes: 1st vs. 2nd presentation), as reported by Li, Miller, and Desimone (1993). From "The representation of stimulus familiarity in anterior inferior temporal cortex," by L. Li, E. Miller, and R. Desimone, 1993, *Journal of Neurophysiology*, 69, p.1925. Copyright 1993 by the American Physiological Society. Permission pending.

Repetition-related increases in neural activity have also been observed in a few other studies, but the circumstances in which they have occurred do not appear to be as general as needed in order to account for priming. Increases have occurred primarily in certain versions of working memory tasks (e.g. when distractors in a DMS task are repeated: Miller & Desimone, 1994; Miller, Erickson, & Desimone, 1996; see Brown & Xiang, 1998 for review) or when stimuli are novel (observed thusfar only in neuroimaging: Henson, Shallice, & Dolan, 2000).

The lack of robust repetition-related activity increases is potentially problematic for distributed connectionist models of priming in light of the data observed by Schall and colleagues (e.g. Hanes & Schall, 1996; Schall, 2001). Using neural recording techniques in monkeys performing an eye movement ("saccade") initiation task, Hanes and Schall (1996) have shown that the initiation of a motor response corresponds to a critical threshold of firing rate in movement-related motor cortex neurons (found in frontal eye field, FEF, in the case of eye movements). In other words, reaction time is determined by the amount of time it takes for motor neuron activity to build up to a fixed threshold (see Figure 1.4). This implies that the motor neurons critical for initiating responses in a priming task must somehow be depolarized more vigorously with stronger input in order to speed up response times.

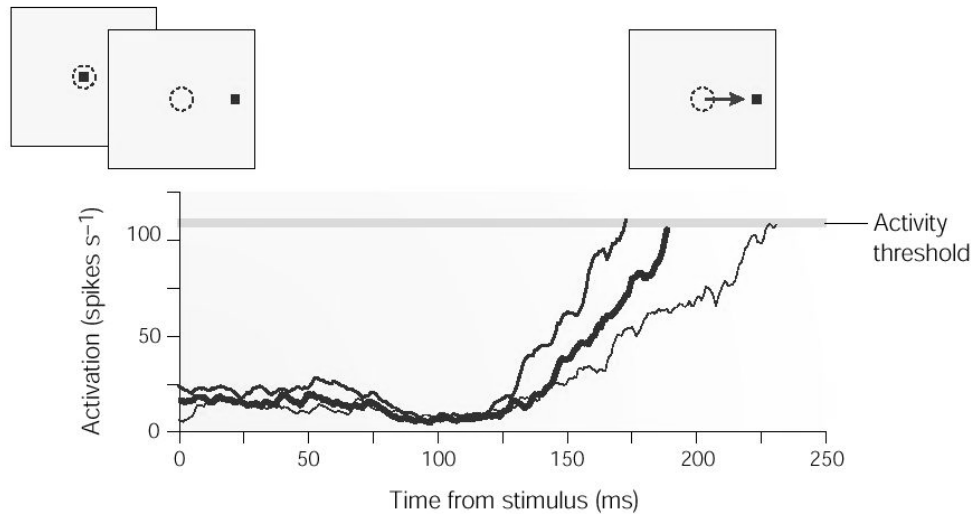


Figure 1.4: Firing rates of neurons in monkey motor cortex (FEF) during the performance of a saccade initiation task, as reported by Schall (2001). Reaction times were reliably associated with the amount of time required for movement-related cells in FEF to reach a fixed firing-rate threshold. From "Neural basis of deciding, choosing, and acting," by J. Schall, 2001, *Nature Reviews Neuroscience*, 2, p. 38. Copyright 2001 by Macmillan Magazines Ltd. Permission pending.

If one assumes that motor cortex neurons representing alternative responses in a given task compete with each other locally through lateral inhibitory, GABAergic synaptic connections as many theorists have argued (e.g. Desimone & Duncan, 1995; Grossberg, 1976; McClelland & Rumelhart, 1981; Usher & McClelland, 2001; Wang, 2002), what is needed to increase firing in critical motor cortex neurons is enhanced excitatory input, decreased lateral inhibitory input, or some combination of both. These conditions might be met if neural activity in sending cortical regions is decreasing overall with practice yet is becoming more selectively tuned because the impact of decreased excitatory support for competing alternative responses (i.e. removal of lateral inhibition) may outweigh small decreases in excitatory support for the correct response. However, increased selectivity alone does not guarantee that the conditions will be met as it is the critical

balance of support for the correct response relative to support for alternative responses that determines the outcome; the increases in selectivity must be of a sufficient degree given the overall activity decreases. To the extent that the mechanisms of lateral inhibition and long-lasting synaptic plasticity are general across cortical brain regions - as preliminary evidence indicates that they are (Bear & Malenka, 1994; Desimone & Duncan, 1995; Kirkwood, Rioult, & Bear, 1996; White, 1989) - one must also ask why firing rate increases would not be observed in each region involved in processing as opposed to only at the point of motor output, especially since learning related decreases are present in these regions. The situation is worsened when one notes that some studies of repetition-related neural changes over many experimental trials have observed both decreased activity and decreased stimulus-selectivity (e.g. Li et al., 1993, p. 1924). Such observations suggest that the greatest decreases in activity are in the neurons critical for generating the correct response, predicting slower buildup of neural activity and slower reaction times.

The findings related to short-term repetition suppression and priming (over the order of seconds) appear even more problematic for existing connectionist accounts. Miller et al. (1993; see also Miller et al., 1991) have shown that repeated visual stimuli separated by several seconds lead to robust activity decreases in inferotemporal cortex of monkeys (down to as little as 40-50% of the initial spiking response over several repetitions). Because the activity decreases tend to be larger for stimuli that initially elicit the strongest spiking responses, the stimulus-selectivity of individual neurons is decreasing overall across repetitions. It is unclear how long-term plasticity mechanisms of LTP/LTD could explain such transient effects; others have also emphasized the ability

of LTP/LTD to enhance stimulus selectivity and "sharpen" neural representations rather than decrease selectivity (e.g. Song, Miller, & Abbott, 2000; van Rossum et al., 2000)⁷. It is important to point out that short-term repetition priming effects in humans occur over these same time scales, even with several intervening stimuli occurring between repetitions (McKone, 1995, 1998; see Figure 1.5). Recent neuroimaging studies involving continued exposure to a stimulus or using repetitions separated by delays of seconds have further established the relation of short-term decreases in neural activity to short-term repetition priming (e.g. Grill-Spector & Malach, 2001).

As I will demonstrate in a series of neural network simulations in this thesis, if one takes into account the generally observed short-term plasticity processes of firing-rate adaptation and synaptic depression (discussed above in **Section 1.1**), the empirical characteristics of short-term repetition suppression are nicely reproduced, including the detailed dynamics of suppression and the finding of decreased selectivity. While it is straightforward to account for the short-term changes in neural activity with adaptation and synaptic depression, the basis of behavioral facilitation is still unclear. When these mechanisms are included within a standard "firing-rate" connectionist neural network that encodes multiple stimuli with competitive interactions, stimulus repetition leads activity to build up slower and at lower levels, leading any fixed criterion of firing rate to be reached more slowly or not at all.

⁷ It is clear from studies involving stimulus repetition following both short and long delays, that activity decreases do not completely recover over longer delays (e.g. Li et al., 1993; Rainer & Miller, 1999). Small, residual activity decreases appear to build up over repetitions spaced by several minutes, consistent with the time course of longer-term plasticity effects such as LTP and LTD. Even in the event that longer-term plasticity mechanisms are expressed at the faster time scale of seconds (note the steady-state reaction time changes in Figure 1.5 after delays of 10-15 seconds), changes of the magnitude required to yield activity decreases of 50% over a few seconds would be expected to yield large interference effects across different stimuli/experiences and would likely undermine the stability of the long-term knowledge base represented across neocortical synapses (e.g. McClelland, McNaughton, & O'Reilly, 1995). See **Section 3.2.5** for further discussion.

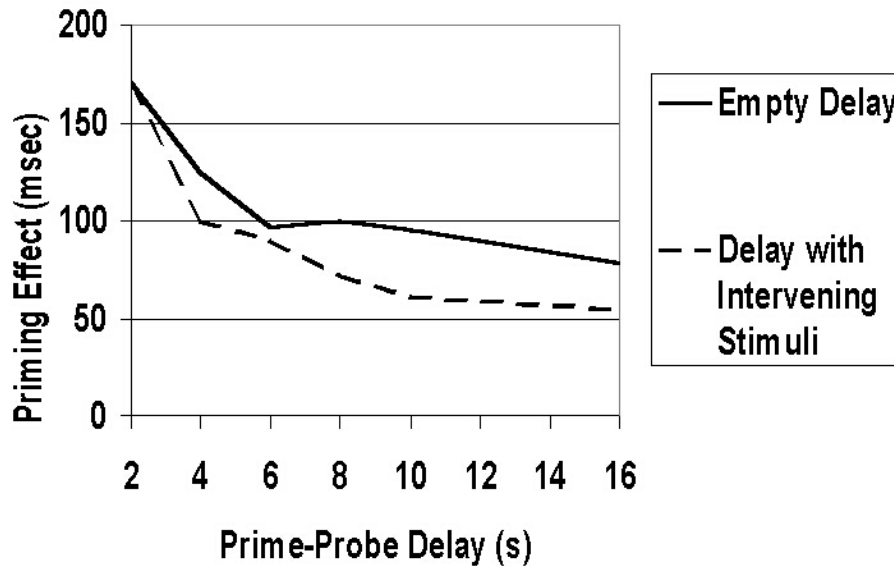


Figure 1.5: Short-term repetition priming effects in human subjects performing a lexical decision task, as reported by McKone (1998). Priming effects show large reductions in magnitude over several seconds, regardless of whether intervening trials are present during the prime-probe delay. From "The decay 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.

Support for this claim can be found in **Appendix A**. As elaborated in the next section, this problem results mainly from the choice of a reduced firing-rate description for neural activity. If one incorporates the spiking property of neural cells, there is a natural way to account for the commonly observed decreases in firing rate with improved behavioral performance.