Cognitive Science 27 (2003) 403–430

COGNITIVE SCIENCE

http://www.elsevier.com/locate/cogsci

# Persistence and accommodation in short-term priming and other perceptual paradigms: temporal segregation through synaptic depression

David E. Huber*, Randall C. O'Reilly

*Department of Psychology, University of Colorado, 345 UCB, Boulder, CO 80309-0345, USA*

Received 1 April 2002; received in revised form 3 December 2002; accepted 20 January 2003

---

## Abstract

Perceptual input changes constantly in an unpredictable fashion, often changing before our somewhat sluggish perceptual systems have adequately processed this input. This can give rise to *source confusion*—how do we know whether a given perceptual activation is due to the current input, or a previous input that had yet to be completely processed? We propose that activity-dependent neural accommodation naturally limits this source confusion by suppressing items once they have been identified. We review behavioral paradigms from different literatures that measure the correlates of persistence and accommodation. Of the various accommodative mechanisms, we focus on synaptic depression, deriving a rate-coded expression that can be used to produce accommodating dynamics in any neural network with real valued activation. We implement this expression in a hierarchical model of perception termed, "a neural mechanism for responding optimally with unknown sources of evidence" (nROUSE). This model can be viewed as a more detailed version of the more abstract ROUSE model of Huber, Shiffrin, Lyle, and Ruys (2001), which produces accommodated levels of feature evidence through an optimal calculation. We apply nROUSE to three short-term priming experiments that manipulated prime duration.

*Keywords:* Associative priming; Repetition priming; Visible persistence; Synaptic depression; Habituation; Perceptual identification; Neural networks; Bayesian; Belief networks; Forced choice; Inhibition of return; Source confusion; Contrast priming; Assimilation priming; Missing dot; Inverse duration; Semantic satiation

---

* Corresponding author. Tel.: +1-303-492-4574; fax: +1-303-492-2967.
*E-mail address:* dhuber@psych.colorado.edu (D.E. Huber).

doi:10.1016/S0364-0213(03)00012-0

## 1. Introduction

One of the most ubiquitous observations in neural processing is that of rapid accommodation with prolonged stimulation. This is seen most clearly in single-cell recording studies in which the rate of neural firing declines from an initial peak, following presentation of a preferred stimulus (e.g., Duysens, Orban, Cremieux, & Maes, 1985). We propose that such rapid accommodation serves an important role in perceptual processing. Specifically, we suggest that accommodation helps avoid the potential hazards of another ubiquitous aspect of neural processing: ongoing persistent neural response beyond removal of the preferred stimulus. The brain is slow to respond, requiring integration over time, which results in perceptual persistence. Accommodation provides a mechanism to clear the system, allowing unobstructed processing of subsequent input. Most importantly, accommodation is driven by activation, and, therefore, items are only suppressed once they have been adequately identified. Therefore, accommodation has no single onset time and is expected to occur at different rates for different regions of the brain and at different rates depending upon experience with different items. In this manner, accommodation is capable of parsing, or segregating, the ongoing stream of constantly changing perceptual stimuli in a relatively automatic manner.

In the first section of this article we discuss the behavioral and neurophysiological correlates of perceptual persistence and neural accommodation. A number of behavioral paradigms demonstrate that short presentations facilitate responses to similar stimuli, whereas longer presentations provide less facilitation, or even inhibit responses to similar stimuli. We interpret these results in terms of persistence, resulting in facilitation for similar or identical items, followed by accommodation, which lessens or reverses this facilitation. With competitive inhibition, the opposite pattern is expected for dissimilar items—when a given item is strongly activated, it will strongly inhibit dissimilar competitors, but when it accommodates, this inhibition will decrease. We propose that these dynamics take place at many levels of processing, from lower sensory areas up to higher-level association areas. Different levels will exhibit different time courses for this biphasic response, with higher levels being typically slower to activate and therefore slower to accommodate. There are a large number of well-specified low-level biphasic phenomena, such as center-surround inhibition, color opponency, light/dark adaptation, etc., which employ similar inhibitory and accommodative dynamics. Our particular implementation may likewise apply in these situations, although they are not the focus of this paper. Instead, our discussion and modeling focus on somewhat higher-level perceptual and cognitive phenomena, which are not typically interpreted in terms of such neural dynamics.

Neurons likely have several accommodative mechanisms operating at different time scales, yet we focus in particular on synaptic depression for its ability to capture neural behavior without appealing to other synaptic mechanisms. Specifically, Tsodyks and Markram (1997) performed a series of neurophysiological studies, and were able to adequately describe the short-term transient coupling of neurons using only synaptic depression. Although our modeling focuses on synaptic depression, we assume that other transient accommodative (and facilitory) mechanisms are likely subsumed in our particular parameterization. We derive a generally applicable "rate-coded" expression for synaptic depression that could be used to modify any of a large number of neural network and dynamic models, allowing greater ability to handle complex sequences of perceptual input.

In the second section, we explore the effects of synaptic depression in the context of a neural network model of word identification called nROUSE ("a neural mechanism for responding optimally with unknown sources of evidence"). In the model, the offsetting components of persistence and accommodation are related to the offsetting components of source confusion and discounting, proposed by Huber et al. (2001) (see also Huber, Shiffrin, Lyle, & Quach, 2002a; Huber, Shiffrin, Quach, & Lyle, 2002b). In their short-term word priming studies, short prime presentations resulted in a bias for repeated or related words, whereas longer presentations resulted in a bias against repeated words. These results were captured with a Bayesian model, ROUSE, containing source confusion (i.e., the assumption that the activation of primed features lingers, becoming mixed with target activation) and an optimal decision process that assigns a discounted level of evidence to primed features. In the online Appendix[1] we present a reformulation of ROUSE in terms of a Bayesian belief network. This reformulation makes is clear that discounting is the same concept as "explaining away" (i.e., the activation of primed features is explained away by the knowledge that the prime may have been the source of activation). Furthermore, the reformulation allows for continuous prime and target activation states, which is an important step for relating ROUSE to the neural network nROUSE model. We demonstrate that synaptic depression in nROUSE produces the same effects as the optimal calculation in ROUSE. Furthermore, we show that nROUSE can capture the data of several experiments manipulating prime duration. Finally, we briefly sketch other recent results that specifically address important assumptions of nROUSE.

Some may find our chosen level of modeling dissatisfying, as we attempt to bridge the gap between low-level neurophysiology and high-level abstract cognitive modeling. Our approach is to employ the minimum number of mechanisms to adequately describe the behavioral phenomena, but, at the same time, make sure that the chosen mechanisms are grounded in accepted neurophysiology. Therefore, we leave out many aspects of neural behavior, and, furthermore, greatly simplify the ones that remain. At the same time, the remaining neurophysiology complicates the abstract functions required to explain human behavior. This is the unavoidable consequence of theorizing across domains of expertise, yet we feel that this is an important endeavor that will ultimately inform and constrain both the neurophysiology and the explanation of cognitive phenomena.

## 2. Biphasic responses in high-level cognitive phenomena

A number of behavioral paradigms are readily interpreted in terms of persistent and accommodating responses. That the paradigms we consider all involve visual processing is largely the byproduct of the authors' areas of expertise, but we feel that these ideas should apply more generally to perceptual processing in other modalities. Our discussion is limited to relatively high-level cognitive phenomena. This is intentional, providing a novel interpretation of these phenomena, which are not usually characterized in terms of automatic neural processes. In contrast, low-level perceptual effects such as center-surround inhibition, color opponency, light/dark adaptation, and many other temporally graded phenomena are typically explained in terms of specific neural dynamics. Our theory may equally apply to these low-level

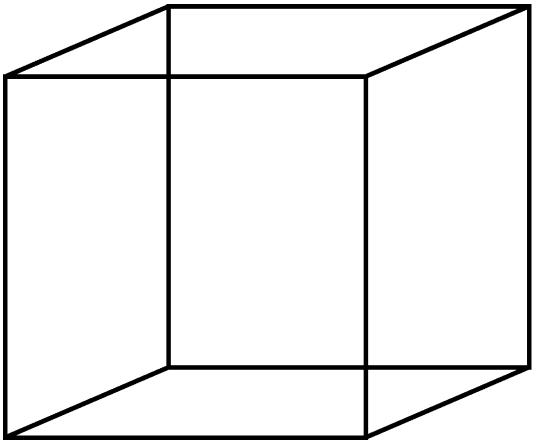

Fig. 1. Necker cube. Either the left or the right face can be seen as the front, and the interpretation will spontaneously alternate between these two views, presumably due to accommodation of the current view.

phenomena, but it is our attempt to explain high-level cognition through the neural dynamics that is unique, and this is our focus at present.

The notion of lateral inhibition is important for explaining the various transient biphasic results considered below (i.e., localized neural areas are limited in their total activation due to non-specific inhibition). Items that are similar or identical will contact the same neural representation and, therefore, persistence will facilitate processing. However, residual activation causes a deficit for dissimilar items due to lateral inhibition (i.e., because something else is active, inhibition is greater, making it more difficult to activate subsequent items).

The phenomenological experience of accommodation, and its effect on lateral inhibition, is found in the spontaneous alternation of ambiguous figures. For example, the Necker cube (Fig. 1) can be seen with either the left or the right face in front, and, furthermore, the perceptual system will naturally flip back and forth between these views. The two views of the Necker cube are placed into competition through lateral inhibition. Initially, one view wins the competition, inhibiting the alternative view. However, once that view is weakened through accommodation, then the alternative view will spontaneously emerge to win the competition. A recent model of the related phenomenon of binocular rivalry, where the two eyes are presented with two different images and perception spontaneously alternates between the two, demonstrates a good fit to the behavioral data using an accommodation mechanism like that used here (Laing & Chow, 2002).

The canonical patterns of persistence, accommodation, and lateral inhibition that we propose are illustrated in Fig. 2, in the context of a simple priming experiment where a prime stimulus is presented for a variable duration, followed by either a direct measure of persistence for the prime, or an indirect measure through immediate presentation of a similar or dissimilar target item. As a function of prime duration, persistence of the prime representation initially increases as the prime accrues greater activation, but with longer durations this reverses, and persistence decreases due to accommodation. For paradigms that measure the indirect effect of a prime on a target item, the relative processing speed of the target is the measure of interest. Because of lateral inhibition, the direction of the biphasic response is opposite for dissimilar target items

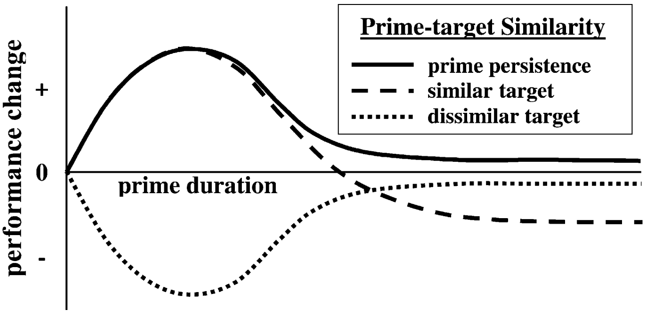

Fig. 2. Hypothetical facilitation and inhibition curves for the effect of a prime stimulus as a function of prime duration. The solid line shows persistence for the prime. As duration increases, persistence also increases, but then decreases due to accommodation. The dotted line shows the effect of the prime upon a subsequently presented dissimilar target item. There are no direct persistence effects upon a dissimilar target, but through lateral inhibition there is a deficit that is related to magnitude of prime persistence. The dashed line shows the effect of the prime upon an identical or similar target item. For short prime durations there is a benefit, due persistence, but at longer prime durations there is a relative deficit, because accommodation makes it difficult to reactivate the shared representation.

(i.e., the extent of target inhibition is directly related to the degree of prime persistence). For a similar or identical target, the situation is more complicated. In this case, persistence gives the target a head start, facilitating activation of the target. In contrast, accommodation results in relatively sluggish activation of the target as compared to a situation in which no prime was presented. For these reasons, short prime durations will facilitate similar targets, but long prime durations can result in a relative deficit for similar targets.

### *2.1. Direct behavioral measures of persistence*

Persistence can be directly measured using neurophysiological methods, but is difficult to assess with a behavioral measure. Nevertheless, there are few techniques that directly measure persistence. One example is Sperling's (1960) partial report paradigm, in which participants report part of a briefly presented display, as directed by a cue appearing after the display. However, this paradigm is limited to short prime durations because longer prime durations allow encoding into long-term memory. Another technique is to require integration between the prime display and a subsequent display. If the display is sufficiently complex, long-term memory is not useful, and longer prime durations are possible. An example of a task requiring integration is the missing dot paradigm (e.g., Coltheart, 1980; Hogben & Di Lollo, 1974). Half of a matrix of dots is presented in an initial display and the other half of the matrix, minus one dot, is presented in a subsequent display. The task is to report the position of the missing dot. This task requires some degree of visible persistence and integration across the two displays. As the duration of the first display increases up to 100 ms, the first display gains stimulus energy, resulting in greater persistence, which improves performance. Beyond 100 ms the opposite occurs and performance decreases, implying a decrease in persistence. This so-called "inverse duration effect" has been attributed to various decay or accommodative mechanisms that accrue with increasing display duration (e.g., Busey & Loftus, 1994; Dixon & Di Lollo, 1994).

Next we consider paradigms from different domains that indirectly measure prime persistence, as realized through the effect of the prime upon the ability to rapidly identify a subsequent stimulus (i.e., the target). Here we see that the effects of prime-target similarity (Fig. 2) and levels of representation (where higher levels activate and thus accommodate more slowly) are important.

### 2.2. *Perceptual priming*

In the domain of 3-D perception, priming has been used to influence the interpretation of ambiguous 2-D targets. In particular, Long, Toppino, and Mondin (1992) had participants view *unambiguous* Necker cube figures for prime durations ranging from 1 to 150 s (ambiguity was removed by rotation and occlusion in separate conditions). For short prime durations, participants initially interpreted a subsequently presented ambiguous Necker cube in accord with the prime presentation, and for long prime durations the opposite was true. These results led the authors to conclude that the priming effects result from neural fatigue for the pre-exposed interpretation. Persistence of the primed view facilitates that view for short prime durations whereas accommodation of the primed view with long prime durations allows the alternative view to emerge.

Leopold, O'Toole, Vetter, and Blanz (2001) also studied so-called "high-level aftereffects," but in the domain of face perception. Using morphing technology, they constructed a face space in which target faces and prime anti-faces lay at opposite ends of a similarity trajectory. Therefore, in this face space, the target face was the most dissimilar, and other potential targets were closer to the prime. The measure of interest was the threshold along the anti-face to target-face trajectory at which the target face was correctly identified (in an initial phase of the experiment, participants learned the four possible target faces). Following 5 s of pre-exposure to the anti-face, they found that this threshold was significantly lowered, such that the target face was more readily identified. Assuming that 5 s was sufficient to induce accommodation, this finding of facilitation is sensible: accommodation to the anti-face made it difficult to activate the alternative responses (which were more similar to the anti-face prime), thereby biasing activation in favor of the target.

Huber et al. (2001) used a two-alternative forced choice procedure (2AFC) and found evidence of both benefits and deficits for repeated words. In their perceptual identification paradigm, a target word was briefly flashed (e.g., around 50 ms), followed by presentation of a backwards pattern mask to ensure elimination of the iconic memory, and finally a choice between two words (Fig. 3). Prior to the target flash, two prime words were presented, allowing priming of the correct target word as well as priming of the incorrect foil word in order to ascertain whether a pattern of costs and benefits occurred with priming (i.e., a bias), or if a perceptual advantage was imparted to the target beyond that occurring for a primed foil. For repetition priming, they found that short (i.e., 500 ms) passively viewed prime durations resulted in a strong preference (i.e., a bias) to choose the word that was just primed. This resulted in better performance when only the target was primed (i.e., the *target* condition) and worse performance when only the foil was primed (i.e., the *foil* condition), as compared to the presentation of unrelated prime words (i.e., the *neither* condition). Remarkably, for longer, actively processed prime viewing (e.g., taking 2,000 ms to give a response in relation to the primes),

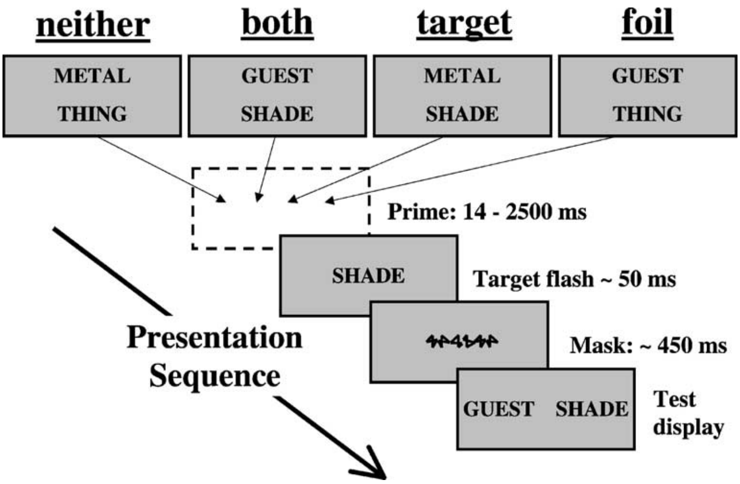

Fig. 3. Perceptual identification with forced-choice testing. This short-term priming paradigm was developed by Huber et al. (2001) and is the basis of the experiments appearing in Figs. 8–10. The particular word examples in Fig. 3 demonstrate repetition priming, although the same four conditions can be tested with other types of priming, such as associative or orthographic similarity between primes and the test display choice words. The neither condition assesses baseline performance. In the both condition, any strategic responding would apply equally to both choice words, and, therefore, this condition provides an unbiased assessment of perceptual benefits or deficits. In the target and foil conditions, only one of the choice words is primed, and these conditions assess whether there is a preference for, or against, primed words.

the pattern of costs and benefits reversed, suggesting a preference *against* choosing repeated words. Subsequent experiments demonstrated that this preference change occurs as the result of prime duration alone, without need of a second task in relation to the primes (Huber et al., 2002b; Huber, in preparation). Persistent processing of the primes explains the preference for repeated words following short prime durations whereas accommodation with longer prime durations explains the switch to a preference against repeated words. In the second section of this article we return to these results and discuss various theoretical interpretations of these and other short-term priming data.

### 2.3. *Semantic priming*

The phenomenological experience of semantic satiation, in which repeatedly speaking aloud the same word results in the loss of meaning for that word, is another example of accommodation. It has proven difficult to quantitatively measure this effect, although the use of priming measures have documented its existence. Specifically, Smith and Klein (1990) had participants speak aloud a category name (e.g., TOOLS) either 3 or 30 times. Subsequently, two target words appeared and participants quickly decided whether the target words were both from the same category (e.g., APPLE and PLUM). When the repeatedly category (e.g., FRUIT) matched the target words, there was a priming facilitation in the form of faster responses for both 3 and 30 repetitions of the category name. However, evidencing semantic satiation, the magnitude of this priming was significantly reduced following 30 as compared to 3 repetitions. Considering that the level of similarity falls short of identity, and that semantic representations may take much longer to fully activate and accommodate, it is in keeping with our theory that the facilitation did not reverse, but rather lessened with accommodation.

### 2.4. *Affective priming*

If persistence and accommodation are truly ubiquitous, then they should occur at all types of processing, including those of social concepts and affective responses. In the field of social cognition, Murphy and Zajonc (1993) investigated short-term evaluative priming as a function of prime duration. In Experiment 1, they presented pictures of angry or happy faces for either 4 or 1,000 ms. This prime presentation was immediately followed by a neutral Chinese ideograph, to which participants gave a liking rating. For the 4 ms prime duration, there was a strong bias affect such that the target ideographs were liked better following happy faces than following angry faces (referred to as *assimilation priming* in the social-cognitive domain). However, for the 1,000 ms prime duration, this bias was completely gone and there was a nearly significant effect in the opposite direction (referred to as *contrast priming* in the social-cognitive domain). The authors interpreted these results in terms of rapid affective access at 4 ms, versus a slower, offsetting, cognitive appraisal at 1,000 ms. Our interpretation also appeals to offsetting components (i.e., persistence vs. accommodation), but we suppose these occur within the same affective representation over the time course of processing.

### 2.5. *Attentional priming*

Next we consider the phenomenon termed "Inhibition of Return" (IOR), which is commonly thought of as an attentional effect, rather than a priming effect. However, if an attentional map is assumed, in which spatial regions of the display are highlighted, then IOR is analogous to the above short-term priming paradigms in that cueing some part of the attentional map affects other parts of the map as dictated by persistence and accommodation.

In IOR, cueing one spatial location of a display results in a biphasic effect on targets presented at that location, dependent upon the Stimulus Onset Asynchrony (SOA) between cue and target (see Taylor & Klein, 1998, for a recent review). On a typical trial, participants view three horizontally aligned boxes, remaining fixated on the center box, and one of the two peripheral boxes is brightened in advance of presenting a target in one of the three boxes. If the brightening cue occurs at the correct location less than 150 ms prior to target onset, targets are responded to more quickly, whereas if the brightening cue occurs at the correct location 300 ms or longer before target onset, targets are responded to less quickly. Throughout the trial, participants do not move their eyes and remain focused on the center box, which is the most likely target location. Differing accounts of these results locate the effect in the attentional system or in the motor response system. In either case, persistence and accommodation could provide the appropriate dynamics, resulting in facilitation for shorter cue-target SOAs and inhibition for longer cue-target SOAs. Indeed, a neural network model of this task exhibits this dynamic based on persistent activations and an accommodation mechanism (O'Reilly & Munakata, 2000).

## 3. Synaptic depression

All of the above results suggest a role for activity-dependent accommodation. In other words, the effect of a first stimulus upon a response or second stimulus is ultimately reduced as a function of increased presentation time and/or SOA between the stimuli. There are several neural mechanisms that can explain activity-dependent accommodation (e.g., calcium-dependent potassium channels, GABA-B inhibitory channels with longer time constants, and synaptic depression). Furthermore, there are also mechanisms for transient synaptic facilitation (e.g., Nelson, Varela, Gibson, & Abbott, 1997). Although a fully realistic biological model should ideally include all of these mechanisms, we have instead focused on the single mechanism of synaptic depression for simplicity, with the idea that the parameterization of this function likely includes the effects of these other mechanisms. This is justified in part from the finding that the activity-dependent effects of coupled neocortical pyramidal cells are adequately characterized using a mathematical instantiation of synaptic depression (Tsodyks & Markram, 1997), and that depression appears to be the dominant form of transient neural plasticity (Nelson et al., 1997).

Synaptic depression occurs when one or more of the resources necessary for synaptics transmission depletes as a result of pre-synaptic activity. With this depletion, pre-synaptic action potentials have a reduced effect, post-synaptically. Candidate resources include neurotransmitter vesicles, calcium (which is necessary for neurotransmitter vesicle release), and activity-dependent post-synaptic de-sensitization. The dynamics of synaptic depression have been studied by patch-clamping two connected neurons and recording the response of the receiver neuron to various trains of sender action potentials. A specific example of such recordings, found in Tsodyks and Markram (1997), appears in Fig. 4, which shows a sender neuron producing a Poisson sequence of action potentials. Despite a stable rate of firing, the initial pre-synaptic action potentials result in a large post-synaptic depolarization, whereas, after about 200 ms, the post-synaptic effect of subsequent action potentials is reduced to a low, asymptotic level.

Much of the work on synaptic depression has focused on the functional characteristics of depressing synapses. For instance, Abbott, Varela, Sen, and Nelson (1997) observed that

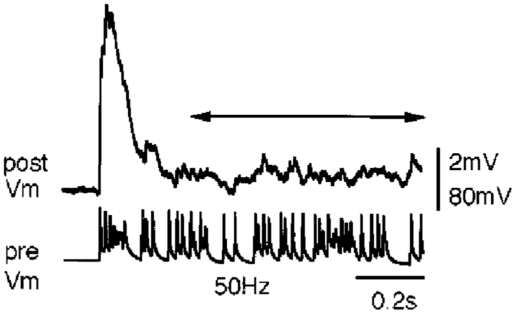

Fig. 4. Electrophysiological recordings obtained by Tsodyks and Markram (1997) with patch clamping of two connected neurons. In this example, the pre-synaptic membrane potential (pre Vm) reveals a random, but relatively consistent train of action potentials. Following a large initial depolarization, the post-synaptic membrane potential (post Vm) rapidly drops to a low asymptotic level. This neural accommodation is assumed to occur due to synaptic depression as dictated by the depletion of resources at the synapse.

depressing synapses serve as a gain control mechanism such that equal changes in pre-synaptic input produce equal post-synaptic changes, regardless of the baseline pre-synaptic firing rate. In another example, Pantic, Torres, and Kappen (2001) simulated integrate and fire neurons, observing that depressing synapses produce enhanced coincidence detection of pre-synaptic action potentials. These same authors have also employed Hopfield networks, demonstrating that synaptic depression allows rapid switching between memory states (Pantic, Torres, & Kappen, 2002) at the expense of reduced storage capacity (Torres, Pantic, & Kappen, 2002). Some researchers have gone beyond these functional observations, suggesting behavioral correlates of synaptic depression. For instance, Manor and Nadim (2001) observed that depressing synapses allow a coupled network to switch between different modes of oscillation. Furthermore, they outlined how such behavior could guide central pattern generators in producing motor output. Chance, Nelson, and Abbott (1998) performed simulations of primary visual cortex and found that synaptic depression enabled bandpass filtering, greater sensitivity to transient input, and other characteristics known to exist in human vision. Also, as mentioned earlier, Laing and Chow (2002) simulated binocular rivalry using synaptic depression mechanisms. Most recently, damage to synaptic depression processes has been proposed as the mechanism underlying an acquired neurophsychological deficit in semantic processing referred to as "access/refractory" (Gotts & Plaut, 2002).

Most simulations of depressing synapses have been conducted with spiking neurons. However, rate-code activations offer many practical advantages in neural simulations (O'Reilly & Munakata, 2000), and we show here that it is possible to derive a rate-coded version of synaptic depression. We begin with the equation found in Chance et al. (1998) for the output amplitude of a spiking neuron:

$$\frac{\mathrm{d}a}{\mathrm{d}t} = R(1-a) + a\ln(\rho_D)\sum_{\mu}\delta(t-t_\mu) \tag{1}$$

where $a$ is the post-synaptic amplitude for a single spike, which can range from 1.0 for a fully "recharged" synapse down to zero, $R$ the rate of amplitude recovery, $\rho_D$ the proportion of resources lost with each action potential, $t_\mu$ the time of a spike, and $t$ is the current time. The first part of this equation provides for exponential recovery toward 1. According to the second term of Eq. (1), amplitude decrease as a function of recent spiking activity (the Dirac delta function, $\delta$, is 1 only when $t = t_\mu$, and is 0 otherwise). Thus, the amplitude is reduced by the value $-a\ln(\rho_D)$ with each action potential. To convert to a rate-coded neuron, we replace the discrete spikes with an average rate of spiking over some temporal window of size, $k$, such that a neuron firing with probability $p$ at each time step will produce the following summation expression:

$$\sum_{\mu}\delta(t-t_\mu) = kp \tag{2}$$

Further simplifying, we combine the constants $k$ and $\ln(\rho_D)$ into a single constant, $D$, for the rate of synaptic depression:

$$D = -k\ln(\rho_D) \tag{3}$$

Finally, we define a rate-coded output variable, $o$, which is equal to the product of firing probability, $p$, and the post-synaptic firing amplitude, $a$:

$$o = pa \tag{4}$$

Substituting Eqs. (2)–(4) into Eq. (1), yields the following rate-coded synaptic depression equation for the change in post-synaptic spike amplitude, $a$:

$$\frac{\mathrm{d}a}{\mathrm{d}t} = R(1 - a) - Do \tag{5}$$

As seen in Eq. (5), the output, $o$, drives synaptic depression (with rate constant $D$) and the extent of depression $(1 - a)$, drives recovery (with rate constant $R$). It is important to note that Eq. (5) does not include any terms for the strength of connection (i.e., no "weight" value) between the pre-synaptic and post-synaptic neurons. As such, this version of synaptic depression can be simulated as a general property of the neuron and does not require separate values for each synapse; in determining the signaling from one neuron to another, the general output value, $o$, is multiplied by the connection strength between the two neurons. However, this is a simplification, and variants of synaptic depression may include the connection strength in the amplitude update equation.

Taylor, Cottrell, and Kristan (2002) recently proposed a similar equation for synaptic depression in rate-coded neurons, applying the result to reciprocally coupled inhibitory networks. In our expression (Eq. (5)), the depleted output, $o$, drives further depletion whereas, in their expression, the non-depleted spike probability, $p$, determined further depletion. For relatively recovered (i.e., $a$ near 1.0) synapses, there is little difference between these two versions of synaptic depression, but, as depletion accrues, spike amplitude can catastrophically drop to zero in their expression, whereas an amplitude of zero is never fully achieved with Eq. (5). We have performed simulations with their expression and found that it is very difficult to produce post-synaptic responses similar to Fig. 4, in which a low, but above zero, asymptote is achieved after the initially large response. In addition, we favor our expression for functional reasons specific to the ROUSE model. In accounting for short-term priming preference reversals, Huber et al. (2001) proposed that the discounting of primed features should result in lower, but above zero levels of evidence in favor of the target. In other words, persistent responses to previous stimuli (primes) should not be completely suppressed because these features may also be part of the current stimulus (the target). Likewise, Eq. (5) provides for a reduction, but not elimination, of activation associated with accommodated (primed) features.

In the second half of this article, we implement Eq. (5) in a neural network model of short-term priming. The success of this implementation is important, yet Eq. (5) stands on its own as a method for introducing synaptic depression into the dynamics of any rate-coded cascaded neural network (e.g., Usher & McClelland, 2001). In traditional neural networks, there is often an activation value, ranging from 0 to 1, produced by a sigmoidal "squashing function." Synaptic depression can be implemented in such networks by assigning this activation value to the spike probability, $p$, in Eq. (5), and modifying unit coupling by multiplying the activation values by the spike amplitude, $a$, of the sending unit.

## 4. Synaptic depression in a neural network model of short-term priming

In this section we apply the synaptic depression instantiation of neural accommodation to short-term word priming data. We model the results of several experiments manipulating prime duration (Huber & O'Reilly, in preparation; Huber et al., 2002b; Huber, in preparation). Huber et al. (2001) found that briefly presented, passively viewed, prime words resulted in a preference for repeated words, whereas actively responded to prime words that were presented for longer, resulted in a preference against repeated words. The ROUSE model captured this difference, assuming a high level of evidence for prime information following short durations and a low, discounted, level of prime evidence following long prime durations. Because different levels of prime evidence are free parameters, ROUSE is incapable of making quantitative, *a priori*, predictions of preference as a function of prime duration. In contrast, the synaptic depression model makes specific time course predictions for accommodation (i.e., discounting). Therefore, prime duration experiments are particularly diagnostic and each experiment motivates additional assumptions found in the synaptic depression model. In addition, we reformulate the ROUSE model in terms of a Bayesian belief net, allowing comparison between the dynamics of synaptic depression and optimal evidence evaluation.

### *4.1. The Bayesian ROUSE model*

The synaptic depression priming model of short-term priming borrows much from the Responding Optimally with Unknown Sources of Evidence (ROUSE) model of Huber et al. (2001), but, in addition, provides a neural mechanism for the discounted level of evidence for primed features. Therefore, we refer to the new model as nROUSE. Before presenting the specifics of nROUSE, we briefly sketch the original ROUSE model.

ROUSE consists of an activation process and an inference process (Fig. 5). In the activation process, features contained in choice words are activated by multiple sources, including the prime and the target for the case of a feature contained in a primed target. Critically, it is assumed that the perceiver cannot directly access these sources, and, therefore, must infer them on the basis of the observed feature activation together with some general knowledge about the preceding perceptual events. This general knowledge is expressed in ROUSE in terms of *a priori* estimates of prime and target source activations, which are assumed to be influenced by various task factors such as prime duration, stimulus similarity, etc. If it is estimated that a feature has been primed, the observed feature activation is "explained away" by the prime, and the posterior probability that the activation is due to the target is reduced (refer to the online Appendix for more details of the Bayesian calculation). This reduction of evidence in the face of priming is referred to as *discounting* in the attempt to respond optimally. For excessive estimates of prime feature activation (i.e., high $\hat{\alpha}$), the inference calculation can reverse the direction of preference. In translating these components into the language of nROUSE, source confusion results from persistent feature activation, and discounting results from accommodation through synaptic depression. Below, we provide a specific example, demonstrating synaptic depression's ability to mimic optimal discounting.

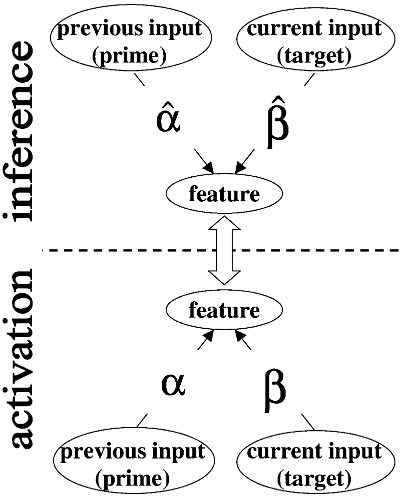

Fig. 5. The Bayesian belief net underlying the ROUSE model of Huber et al. (2001). A primed target feature is independently activated by the prime (with probability $\alpha$), the target (with probability $\beta$), or noise (with probability $\gamma$, which is not shown as a separate source, but factors into the conditional probabilities as seen in the online Appendix). An inference process determines the posterior probability that the target could have been a source of activation, given the estimated activation probabilities ($\hat{\alpha}$, $\hat{\beta}$, and $\hat{\gamma}$). The posterior probability of the target is reduced (discounted) when it is known that the prime is a potential source. In other words, the observed activation of the feature is "explained away" by the prime, reducing the posterior probability of the target. In relating ROUSE to synaptic depression, the prime plays the role of residual activation from previous inputs and the target plays the role of activation from a current input.

### 4.2. The neural ROUSE model: nROUSE

nROUSE is based on a simplified *point neuron* model that captures important aspects of the temporal integration of information in real neurons, while simplifying their geometry to a single mathematical point (O'Reilly & Munakata, 2000). These point neurons have a simulated membrane potential, $v$, that integrates excitatory, inhibitory, and leak currents as dictated by the electrical and diffusion forces operating in real neurons. The general form of this membrane potential update equation is

$$\frac{\mathrm{d}v(t)}{\mathrm{d}t} = \tau \sum_{c} g_c(t)(E_c - v(t)) \tag{6}$$

where $g_c(t)$ and $E_c$ represent the conductance and reversal potential, respectively, of channel $c$, and the sum is over the three channels of excitation, inhibition, and leak. $\tau$ is a time constant reflecting capacitive currents, among other things. The conductance term for excitation, $g_\mathrm{e}(t)$, as elaborated below, reflects the excitatory synaptic inputs from other neurons (often called the net input in neural network models), and the inhibitory conductance, $g_\mathrm{i}(t)$ similarly reflects inhibitory synaptic inputs. The leak conductance, $g_\mathrm{l}(t)$, is a constant.

The probability of spiking, computed as a real-valued number (i.e., a rate code) instead of discrete spikes, is related to the difference between membrane potential and firing threshold, $\theta$, and is zero for membrane potentials at or below threshold. Substituting this spike probability in Eq. (4) results in Eq. (7) for the output, *o*, of a simulated neuron in nROUSE. It is useful to think of this rate code as representing the average response of a population of spiking neurons all of which have similar inputs and outputs:

$$o = (v - \Theta)a \tag{7}$$

We do not claim that the difference between the membrane and firing threshold is literally equal to the probability of spiking, but rather that the two can be related with some constant of proportionality. This spiking probability is the equivalent of activation in traditional neural network theories, as typically determined with a sigmoidal squashing function. However, we do not utilize such a squashing function, instead, allowing the non-linearities inherent in Eq. (6) to saturate activation in an appropriate manner (e.g., Eq. (8) presents a specific instantiation of Eq. (6), which gradually saturates at membrane potential equal to 1.0).

Coupling together Eqs. (5)–(7), the right-hand panel of Fig. 6 shows the response of a simulated unit, with synaptic depression, using parameters consistent with visual layer of the perceptual hierarchy explained below (the figure caption reports the parameter values). In this

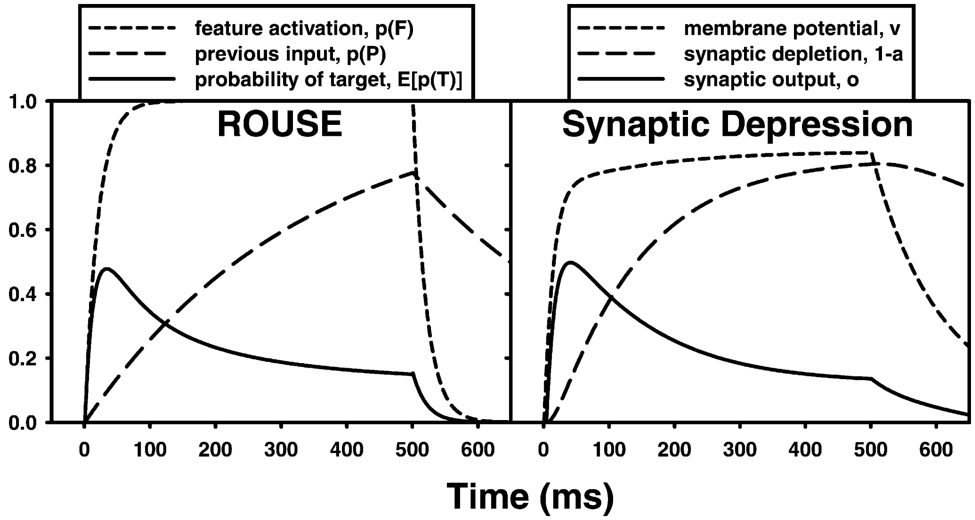

Fig. 6. Comparison between the ROUSE model and synaptic depression. Synaptic output (solid line in right-hand panel) is analogous to the expected posterior probability that feature activation is due to a current input (solid line in left-hand panel), and both are similar to the post-synaptic membrane potential seen in Fig. 4. For synaptic depression, the n-shaped synaptic output (i.e., post-synaptic membrane potential) results from multiplying spike amplitude ($a$), by the rated-coded activation ($v - \theta$). As membrane potential quickly rises, this produces an increase in output, but later, as spike amplitude depletes, output decreases. In ROUSE, the probability of feature activation plays the role of membrane potential (dotted lines in both figures) and the probability of previous input plays the role of synaptic depletion (dashed lines in both figures). Dynamics were added to ROUSE by activating the feature at a fast rate (exponential constant = 0.05) for 500 ms, followed by a deactivation at that same fast rate. The probability that a previous input may have been a source of activation was activated and deactivated at a slow rate (exponential constant = 0.003). ROUSE parameter values were: $T_o = 0.12$, $\hat{\alpha} = 1.0$, $\hat{\beta} = 1.0$, and $\hat{\gamma} = 0$. Synaptic depression parameter values were: $I = 0.3$, $\theta = 0.15$, $L = 0.15$, $D = 0.2$, $R = 0.03$, and $S = 0.05$.

example, a single unit was driven with an excitatory input of 1.0, starting at 0 ms, and then the input was removed at 500 ms. There was no feedback and no other sources of inhibition except for self-inhibition. As seen in the figure, the membrane potential, $v$, quickly reached an asymptotic level, then fell towards zero with stimulus offset, as dictated by the leak conductance, $L$, self-inhibition, $I$, and the rate of integration, $S_V$ (see Eq. (8) for the relation between these constants and Eq. (6)). Synaptic depletion (i.e., $1 - a$) was initially zero, and slowly rose to an asymptotic level as synaptic resources ($a$) were consumed. At stimulus offset, synaptic depletion dropped back to zero as synaptic resources were replaced. Note that the simulated synaptic output is very similar to the post-synaptic membrane potential recorded by Tsodyks and Markram (1997) in Fig. 4.

The left-hand panel of Fig. 6 shows the analog of membrane potential, synaptic depletion, and post-synaptic output, as realized within the Bayesian ROUSE model (the figure caption reports the parameter values used). In order to accomplish this, we reformulated ROUSE as a Bayesian belief network (see the online Appendix), allowing continuous probabilities of feature activation, $p(F)$, which is similar to membrane potential, and continuous probabilities of previous input, $p(P)$, which is similar to synaptic depletion. As seen in Fig. 5, the sources of feature activation in the original ROUSE model were restricted to the target ($T$), primes ($P$), and constant noise activation (which is not shown as a separate source, but instead factors into the conditional probabilities). Relating ROUSE to synaptic depression, prime activation is the equivalent of residual activation from a previous input and target activation is the equivalent of current input from a new item. The inference process explains away activation that may be residual from a previous presentation, allowing accurate estimation of new input. The expected posterior probability of the target as a source, $E[p(T)]$, is this estimation, and, as seen in the figure, is similar to post-synaptic output (see the online Appendix for the derivation of this expression). ROUSE does not contain activation dynamics, so we provided simple exponential activation in order to drive ROUSE in a manner similar to the neural model. The ability of synaptic depression to mimic the optimal inference calculation in ROUSE suggests that the transition between pre-synaptic firing rate and post-synaptic output, as mediated by synaptic depression, may have evolved to calculate the posterior probability that current input is due to novel presentations, rather than residuals from previous presentations.

Besides introducing activation dynamics through integration and synaptic depression, the nROUSE model specifies the representation, placing units into a perceptual hierarchy such that higher levels are more abstract, and lower levels are more closely tied to the perceptual input. In particular, for visual word identification, nROUSE assumes the three processing levels appearing in Fig. 7: visual, orthographic, and lexical-semantic. Units in the visual layer receive an input of 1.0 when the preferred visual stimulus is presented at the corresponding spatial location, but otherwise receive no input. The spatial locations (e.g., the three boxes across the bottom of Fig. 7) include units for all possible visual features at a given location. The visual level has excitatory bottom-up projections to a general pool of orthographic (and possibly phonemic) units that process orthography regardless of spatial location (nevertheless, orthography is sensitive to the relative position of letters within a word). In theory, we could include top-down excitatory feedback between orthographic and visual units, but this was not needed in the current application, and can be problematic, causing the model to "hallucinate" non-existent visual features. Orthographic units have excitatory bottom-up projections to a

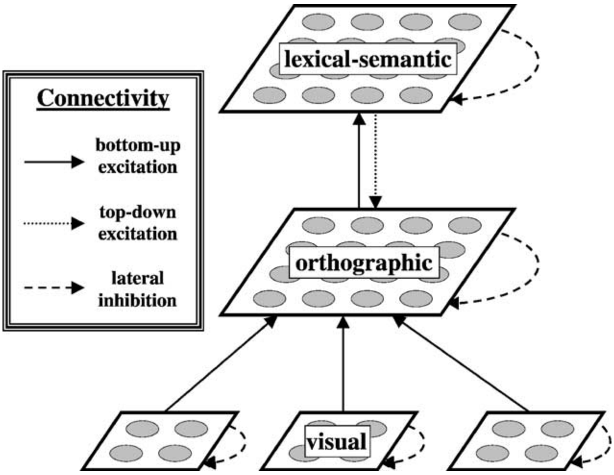

Fig. 7. A perceptual hierarchy for word reading, in the neural mechanism for responding optimally with unknown sources of evidence (nROUSE) model. The bottom level includes visual features for each spatial region, the middle level is a general pool of orthographic features, and the top level is a general pool of lexical-semantic features. Connections between levels in both directions are excitatory. Connections within each square are inhibitory, with every unit connecting with every other unit, including self-inhibition. This "lateral inhibition" limits the amount of total activation within a level and sets up winner-take-all attractor dynamics. See the text for further discussion and details.

general pool of lexical-semantic units. In turn, the lexical-semantic units have excitatory top-down projections (i.e., feedback) to the corresponding orthography. This recurrence sets up "attractor dynamics" between orthographic and lexical processing such that the system can gradually settle into the appropriate state given the orthographic and lexical constraints.

Crucial to the development of attractor, dynamics is an assumption of lateral inhibition within each patch of units (i.e., within each box in Fig. 7), which is produced in the brain by local inhibitory interneurons. In the model, we simplify the situation, using direct inhibitory connections among all the units within a patch (including self-inhibition). This results in a winner-take-all dynamic in which the entire network gradually settles upon a presented word.

Incorporating these architectural properties, the complete membrane potential ($v$) update equation for each millisecond of time, for a particular unit $i$, at level $n$ in the perceptual hierarchy, is

$$\frac{\Delta v_i^n(t)}{S_n} = (1 - v_i^n)\left\{\sum_{\forall j} w_{ij} o_j^{n-1} + F \sum_{\forall k} w_{ik} o_k^{n+1}\right\} - v_i^n \left\{L + I \sum_{\forall l} o_l^n\right\} \tag{8}$$

The bottom-up (i.e., connections from the $n-1$ level) and top-down (i.e., connections from the $n+1$ level) excitatory inputs combine to drive the membrane potential towards the excitatory reversal potential ($E_e$) of 1.0 (i.e., the first term in curly brackets, multiplied by $1 - v$). The sum of the bottom-up connections is given a default strength of 1.0, whereas the sum of the top-down feedback, from lexical-semantic units, is given the strength value ($0 < F < 1$). In opposition to the excitatory input, lateral inhibition (i.e., the summation over the $l$ units within level $n$) is given a strength of $I$ and combines with a fixed term for leak, $L$. Together, these

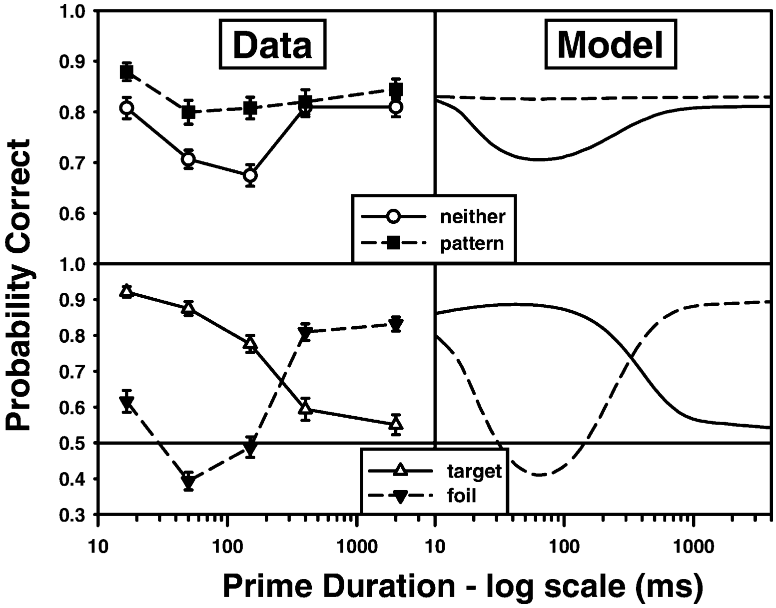

Fig. 8. Huber (in preparation) repetition priming experiment. Only one prime was presented on each trial, appearing directly above and below the midline, with no space between. In the pattern condition, a pattern mask, rather than a prime word, was presented in the above and below positions. Prime presentation durations: 17, 50, 150, 400, and 2,000 ms. Best-fitting nROUSE parameters: $S_V = 0.054$, $S_O = 0.046$, $S_L = 0.015$, $D = 0.324$, $R = 0.022$, and $N = 0.036$.

inputs drive the membrane potential towards the reversal potential of $E_\mathrm{i} = E_\mathrm{l} = 0$ (i.e., the second term in curly brackets, multiplied by $-v$). We assume that each level in the perceptual hierarchy may integrate at a different speed, as dictated by the parameter, $S_n$ (i.e., there are three different $S$ parameters for the three levels of the hierarchy). In general, as seen in the best-fitting parameters reported in Figs. 8–10, higher levels of the hierarchy integrate more slowly (i.e., have smaller $S$ values) than lower levels. This is sensible considering that visual input changes frequently (e.g., across different saccades), whereas high-level concepts (e.g., the meaning of a word), persist for seconds or longer. For a given level, the same rate of integration, $\mathrm{S}_n$, is used in the spike amplitude update, giving rise to Eq. (9) (we assume the same depletion, $D$, and recovery, $R$, strengths for all three levels):

$$\frac{\Delta a_i^n(t)}{S_n} = R(1 - a_i^n) - Do_i^n \tag{9}$$

The weight magnitudes in Eq. (8), are assumed to be 1.0 for learned connections and 0 for unknown, or inappropriate connections. This simplifying assumption is adequate for the studies under consideration because stimulus materials were always drawn from the same class. However, if words in one condition had been of higher natural language frequency than words in another condition, then it would be necessary to either allow different free parameter for the connection weights in each case or to actually simulate learning for the different classes of words. In addition, we assume symmetric weights such that the same value connecting two

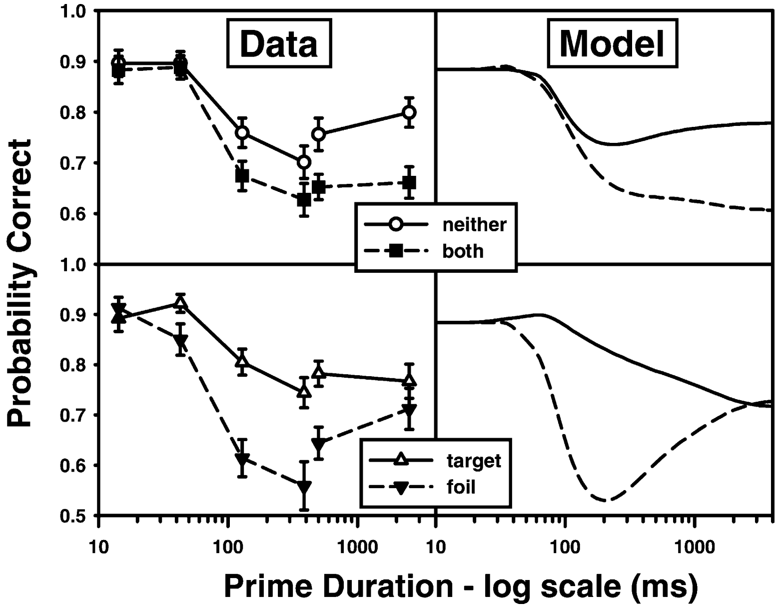

Fig. 9. Huber and O'Reilly (in preparation) for the four shorter prime durations, and Huber et al. (2002b) for the two longer prime duration. Both experiments were repetition priming experiments with two primes presented on every trial, as seen in Fig. 3. In order to capture the difference between a centrally presented prime (i.e., the results of Figs. 8 and 10) and these experiments, which used more peripheral primes, a prime salience parameter (PS $< 1$) was used to scale down the visual input for the primes. Prime presentation durations: 14, 43, 128, and 386 ms (Huber & O'Reilly); and 500 and 2,500 ms (Huber et al.). Best-fitting nROUSE parameters: $S_V = 0.034$, $S_O = 0.075$, $S_L = 0.015$, $D = 0.159$, $R = 0.055$, $N = 0.031$, and PS $= 0.266$.

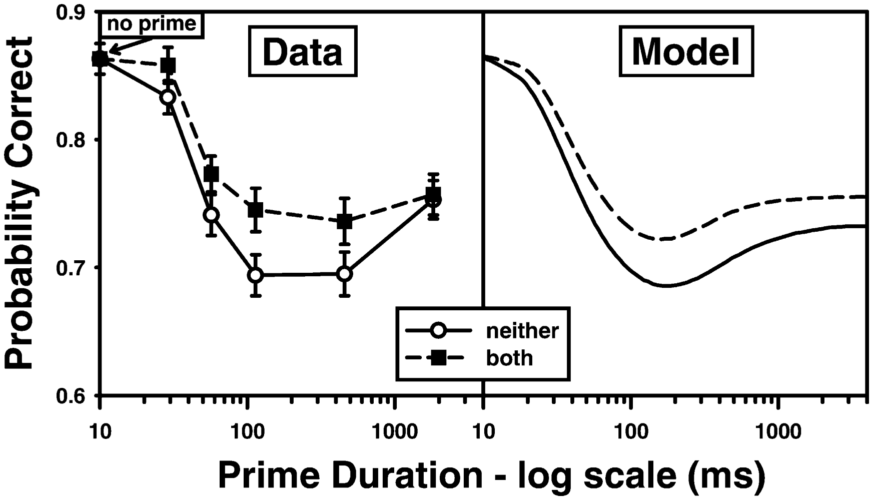

Fig. 10. Huber and O'Reilly (in preparation) associative priming experiment with a single, centrally located prime word. Lexical-semantic similarity was modeled with a parameter for the proportion of shared lexical-semantic features (PL $< 1$). Prime presentation durations: 0, 29, 57, 114, 457, and 1,829 ms. Best-fitting nROUSE parameters: $S_V = 0.046$, $S_O = 0.056$, $S_L = 0.015$, $D = 0.102$, $R = 0.072$, $N = 0.022$, and PL $= 0.267$.

units is used for both bottom-up and top-down feedback (although feedback is scaled by the value, $F < 1$, and there is no feedback to the visual layer; see O'Reilly, 1998; O'Reilly & Munakata, 2000 for discussion of the multiple motivations for such symmetric connectivity).

### *4.2.1. Representations in nROUSE*

Next, we consider the representational scheme used in nROUSE to simulate the behavioral priming experiments. In principle, we imagine that there are distributed representations at each level of the model. However, actually constructing these representations for the corpora of words used in the experiments would result in a very large model that would render precise parameter fitting nearly impossible. Therefore, we employ a simplified localist instantiation of a distributed representation, such that a single unit was used to represent all the features that a given word has in common with other word(s), and additional units were used to represent the features unique to each word. These units were then weighted to reflect the relative extent of similarity or difference between the words. For example, in the case of repetition priming with dissimilar choice words (i.e., having no overlapping features), it was only necessary to simulate one unit for each word at the orthographic and lexical-semantic levels. The visual input required one unit at each spatial location (see Fig. 3) where that word might appear. To see how similar words were represented, consider the word "CAT" priming the choice word "DOG." These words are completely different at the visual and orthographic levels and therefore separate units would be used for each word at these levels. However, cats and dogs are both pets and there is some degree of overlap at the lexical-semantic level. This would be captured with three aggregate units: one for the lexical-semantic features shared between cat and dog (i.e., cat-dog features), one for the cat-not-dog features, and a third for the dog-not-cat features. The output of the cat-dog unit would be weighted by a similarity parameter, PL (i.e., the proportion of shared features at the lexical level), and the cat-not-dog and dog-not-cat units would be each weighted by $1 - \mathrm{PL}$.

All simulations mimicked the exact spatial and temporal characteristics of the experimental display sequences. Pattern masks were assumed to involve as many visual features as words. Therefore, a separate pattern mask visual unit, also of size 1.0, received input at times appropriate to the pattern mask presentation. Because the pattern mask was assumed to be dissimilar from the orthography of any known word, the only effect of the pattern mask was inhibition of the target flash visual unit (i.e., the pattern mask more rapidly "turned off" the visual features of the target). At the visual level, masking is specific to particular spatial regions, as dictated by lateral inhibition within a patch (higher level, "meta-masking" effects can nevertheless occur in the orthographic and lexical levels).

### *4.2.2. Forced-choice responding and reaction time in nROUSE*

In the 2AFC short-term priming paradigm, the participant must choose one of the two choice words. Choice responding in nROUSE is based on the neural activation from presenting these choice words as it propagates through the network in parallel for both words. The only manner in which the target flash affects behavior is via residual target activation (we assume that explicit identification of the target flash does not occur on most trials, due to the threshold-level presentation of the target). Through simulation, we found that a horse-race decision rule, in which the first choice word to reach peak activation is chosen, provides a good indication of residual target activation, whereas other logical alternatives do not. Simply reading out the residual target activation is not appropriate considering that the system does know the choice alternative in advance. Using peak magnitude in response to the choice words does not index residual activation—peak magnitude is greatest for a novel word, whereas previously viewed words have lower peaks due to synaptic depression. Essentially, peak magnitude gauges the extent of depleted resources (which word was seen longer), instead of the degree of residual activation (which word was seen more recently). In contrast, the time-to-peak horse-race decision rule measures residual target activation because this residual activation gives the target item a head start.

As a result of the horse race, nROUSE predicts reaction times as well as accuracy, with a tight coupling between the two. Specific predictions of the decision rule were tested by Huber (in preparation), and are discussed later in the "Future work with nROUSE" section. Residual prime activation (i.e., persistence) speeds the response to a primed word, but synaptic depression (i.e., accommodation) makes the system sluggish to respond to primed words. These factors offset one another to various degrees, causing a preference for or against primed words in a well-specified manner as a function of prime duration.

One concern with the decision rule is reliability in the face of stochastic variability. If synaptic output varies from moment to moment, it is difficult to determine the true peak in light of the many transient false peaks. Although the reported simulations do not employ any noise processes, we have investigated this issue in other simulations. A robust "peak detector" can be constructed from neurons with different rates of integration. In particular, slow integrators accumulate a "go" signal for each of the choice alternatives. However, inhibition from fast integrators suppresses these go signals (i.e., a "no-go" signal). Only release from inhibition, as occurs on the downside of a peak, allows the go signal through to the motor system. The rate of integration of the fast integrator can be modified, allowing robust detection in the face of noise. This modification slows down transmission of the go signal, adding a fixed delay to all conditions, but does not qualitatively change accuracy as a function of priming condition. Such a "go/no-go," response system is known to exist in the basal ganglia, which has been simulated using the same point-neurons employed in nROUSE (Frank, Loughry, & O'Reilly, 2001).

For the reported simulations, we do not assume a specific noise process and instead use a generic noise level, $N$, which reduces performance as a function of the difference between the lexical peak time of the target and the lexical peak time of the foil. The particular function used is the logistic function appearing in Eq. (10), which yields the proportion correct, $p(c)$, given the peak time of the target lexical representation, $T_{\mathrm{T}}$, as compared to the peak time of the foil lexical representation, $T_{\mathrm{F}}$, expressed in milliseconds:

$$p(c) = \frac{\mathrm{e}^{N(T_{\mathrm{F}}-T_{\mathrm{T}})}}{1+\mathrm{e}^{N(T_{\mathrm{F}}-T_{\mathrm{T}})}} \tag{10}$$

### 4.3. Empirical results and nROUSE assumptions

Next, we apply the nROUSE model to several short-term priming experiments that varied prime duration. Beyond synaptic depression, different experiments motivate different components of the nROUSE model (i.e., levels of processing, lateral inhibition, threshold, and feedback). To facilitate comparison of parameters across the experiments, we fixed several parameters to specific "default" values across all experiments, while allowing others to freely vary to capture the data. Specifically, we fixed the "structural" parameters, while allowing the "dynamic" parameters to vary. These fixed structural parameters were: inhibition, $I = 0.3$, feedback, $F = 0.25$, threshold, $\theta = 0.15$, and leak, $L = 0.15$. The freely varying dynamic parameters were: rate of visual integration, $S_V$, rate of orthographic integration, $S_O$, rate of lexical-semantic integration, $S_L$, depletion, $D$, and recovery, $R$. In addition, different performance levels occurred in each experiment, and a noise parameter, $N$ (see Eq. (10)), was used to capture these differences.

### *4.3.1. Repetition priming with a single prime—Huber (in preparation)*

The word repetition priming experiment of Huber (in preparation) is ideally suited to test the core nROUSE proposal of accommodation through synaptic depression. This experiment varied prime duration from 17 to 2,000 ms (see Fig. 8). Unlike the paradigmatic example seen in Fig. 3, only one prime word was presented, appearing simultaneously in the above and below locations. Furthermore, there was no vertical space between the above/below locations, ensuring that participants directly foveated the prime. Because there was only one prime, the both-primed condition could not be tested. To test the levels of processing assumption, one condition presented a pattern mask, rather than a prime word.

As seen in the lower left panel of Fig. 8, there were dramatic preference changes as a function of prime duration. For durations less than 200 ms, there was a strong tendency to choose the repeated (primed) word such that performance in the target condition was better than performance in the foil condition. This preference peaked with a 50 ms prime presentation diving the foil condition significantly below chance. The preference switched at 400 ms, resulting in a preference against choosing repeated words, such that performance in the target condition was worse than performance in the foil condition. As seen in the lower right panel, the nROUSE model captured this highly non-linear pattern through persistence, resulting in positive preference for shorter prime durations, and the offsetting effect of accommodation through synaptic depression, resulting in a negative preference for longer prime durations (see the figure caption for the best-fitting parameter values).

The upper panels of Fig. 8 test the levels of processing and lateral inhibition assumptions of the nROUSE model. In these conditions, preference plays no role because neither choice word is primed. However, there are still indirect, interference effects of the prime due to inhibition from lingering prime activation. nROUSE predicts "u-shaped" interference effects because of its "n-shaped" synaptic output function (see Fig. 6). The prime acts as a forward mask on the target flash, with the strength of masking dependent upon the degree of persistence and accommodation of the prime and the types of features contained in the prime. In the case of a pattern mask prime (i.e., the *pattern* condition), there is only interference at the visual level, whereas an unrelated prime word (i.e., the *neither* condition) also includes orthographic and lexical-semantic features, and therefore greater interference. These predictions are confirmed in the data of the upper left hand panel of Fig. 8. Pattern performance was better than neither performance for all prime durations. Besides confirming these qualitative predictions, nROUSE was able to quantitatively fit these data, as shown in the upper right hand panel of Fig. 8. The best-fitting parameters produced an almost indiscernible amount of u-shaped interference in the pattern condition due to very little persistence at the visual level. To capture the small u-shaped effect for the pattern condition seen in the behavioral data, it would be necessary to give the pattern mask some degree of higher representation at the orthographic or lexical-semantic levels.

### *4.3.2. Repetition priming with two primes—Huber and O'Reilly (in preparation) and Huber et al. (2002b)*

Because the repetition priming experiment of Huber (in preparation) used only a single prime word, it was not possible to include a both-primed condition. Nevertheless, such a condition is of interest from the standpoint of unbiased perceptual benefits and/or deficits—if performance changes when both choice words are primed, this change cannot be due to bias effects, which affect both choice words equally. The repetition priming experiments of Huber et al. (2001) used two prime words and always observed deficits in the both primed condition. However, those experiment only looked at 500 ms or longer prime durations. Next, we consider the repetition priming experiment of Huber and O'Reilly (in preparation), which used shorter prime durations. Fig. 9 presents this experiment (i.e., the four shorter prime durations) combined with the repetition prime duration experiment of Huber et al. (2002b), which tested two longer prime durations using an identical paradigm. The sequence of events for these experiments is the same as shown in Fig. 3.

As seen in the lower left panel of Fig. 9, there were preference changes with prime duration, although these changes were not as large, and were somewhat delayed, compared with the single prime repetition experiment (i.e., compared to Fig. 8). A preference for repeated words emerged (i.e., target better than foil), reaching its peak around 400 ms, followed by diminishing preference, but not a reversal (i.e., at 2,500 ms, target and foil performance were equal). In other words, the rate of preference change was relatively slow with two more peripherally presented prime words (Fig. 9) as compared with a single centrally presented prime word (Fig. 8). With primes presented more peripherally, participants remain focused on the blank area between the primes. We modeled this through a prime salience parameter, PS, which scaled the visual input of the prime words (PS set to 1.0 in Fig. 8). With a lower level of visual input, the primes asymptote at a lower level of output, and, therefore, accommodation through synaptic depression occurs more slowly. With this modification, nROUSE captured the observed preference changes using very similar parameters to those used in Fig. 8 (see the figure captions for best-fitting parameter values). Recently, Weidemann, Huber, and Shiffrin (submitted for publication) demonstrated that both vertical separation (i.e., more peripheral prime presentation), and the number of primes are important factors in determining these preference effects.

The upper panel of Fig. 9 shows the neither and both conditions, as a function of prime duration. As with the single prime experiment, a u-shaped prime interference pattern is seen for the neither condition. As before, nROUSE captured this through increasing inhibition due to increasing persistence, followed by decreasing inhibition due to decreasing persistence (i.e., increasing synaptic depression). In addition, both-primed deficits gradually accrue as a function of prime duration. Due to lexical feedback, nROUSE produces both-primed benefits, as seen in the next experiment. However, synaptic depression at the orthographic level offsets the benefits of feedback, causing both-primed deficits for repetition priming. Synaptic depression of orthographic units limits the amount information gained from the brief flash of the target word. Because these deficits result from synaptic depression, they are seen to monotonically grow as function of prime duration.

Finally, the shorter prime durations in Fig. 9 justify the use of a threshold (i.e., Eq. (7)) in relating the membrane potential to spike probability (there is ample neurophysiological justification for this assumption as well). At 14 ms there was no difference between any of the conditions, and, at 43 ms, there was only a small difference between the target and foil conditions. Because of the reduced rate of prime processing due to the reduced prime salience (see the discussion above), nearly 40 ms of prime processing was required in nROUSE before the threshold was crossed for the orthographic representation of the primes. This thresholding explains the observed lack of priming effects for very short prime durations.

#### *4.3.3. Associative priming with a single prime—Huber and O'Reilly (in preparation)*

The first repetition priming experiment demonstrated a need for different levels of processing in order to handle different levels of forward masking. However, none of the experiments discussed so far motivate the assumption of feedback between levels. The associative priming experiment of Huber and O'Reilly (in preparation) used a single, centrally presented, prime word that was either unrelated to the choice words (i.e., the neither condition) or was associatively related to both choice words (e.g., GRAVY priming the choice words SAUCE and TRAIN). Performance in this both condition was slightly better than the neither condition at nearly all prime durations (see Fig. 10). This perceptual benefit implies some sort of interaction between the prime and the amount of information gained (or retained) from the target flash. Feedback between the lexical-semantic level and orthography provides the necessary interaction to produce both-primed benefits. In the case of repetition priming, as discussed above, synaptic depression of orthographic units counteracts the beneficial effect of feedback. However, the associatively related words used by Huber and O'Reilly (in preparation) were orthographically dissimilar, allowing perceptual benefits to emerge.

As seen in Fig. 10, nROUSE captures both the u-shaped word interference in the neither condition, as well as the associative both-primed benefit, that reaches its peak at the duration of greatest interference (i.e., the prime duration producing the greatest persistence and interference also produces the greatest amount of lexical feedback, for semantically related words).

## 5. General discussion

A variety of behavioral paradigms demonstrate that stimuli persist for some period of time past their removal. Presumably, this persistence is the natural byproduct of the need to integrate information over time. Experiments manipulating stimulus duration find that persistence increases, up to a point, and then decreases with even longer presentations. We attribute this reduction to neural accommodation. Specifically, we propose that accommodation is a useful mechanism, serving to limit persistence of previous stimuli and allow for relatively unobstructed processing of subsequent stimuli. In this manner, accommodation segregates the continual flow of perceptual input such that items are suppressed once they are identified.

Several neural mechanisms produce activity-dependent accommodation, although we focused on synaptic depression for its ability to encompass other activity-dependent synaptic efficacy effects. Synaptic depression arises from the depletion of resources at the synapse with each additional action potential. Electrophysiological experiments have mapped out the time course of synaptic depression, allowing mathematical specification of the underlying dynamics. We derived a rate-coded version of these dynamics that can be applied to any neural network.

Support for accommodation through synaptic depression was found by incorporating synaptic depression into a rate-coded neural network model of short-term word priming. The model consists of a perceptual hierarchy, with each level integrating at a different rate, such that identification, and accommodation, occur at differential speeds at different levels of processing. The model was successfully applied to three different experiments in which a preference for or against primed words was produced by varying prime duration. In addition, perceptual benefits/deficits and masking/interference phenomena were explained.

### *5.1. Similarities and differences between ROUSE and nROUSE*

Both the ROUSE model of Huber et al. (2001), and the nROUSE model, presented here, reduce the influence of previously presented stimuli through the discounting of features. Given a system in which external input contacts the same internal representations and it is unknown by which sources a feature has become active (i.e., source confusion), both models specify appropriate levels of reduced evidence for previously presented features. However, the implementation of the two models is vastly different. Nevertheless, these implementation differences are a strength, with each model informing the other. In ROUSE, feature evidence is the result of an optimal calculation: For a previously presented feature, a discounted level of evidence is assigned according to the estimated activation probabilities. In contrast, feature discounting in nROUSE results from the relatively automatic process of neural accommodation through synaptic depression. In combination, the models suggest there is a need to reduce persistent feature activation from the standpoint of optimal processing and that synaptic depression may have evolved as a means to satisfy this need.

Besides allowing such a "neural to functional" mapping, nROUSE was developed for its ability to handle aspects of the data that were either unknown, or problematic, for the original ROUSE model. In particular, because nROUSE is a dynamic model in which processing is directly related to the sequence of events, it makes specific predictions as a function of stimulus duration. In contrast, different free parameters are necessary for different durations in the original ROUSE model. Therefore, nROUSE requires fewer parameters to handle a range of prime durations, although this reduction comes at the cost of increased complexity (see Pitt, Myung, & Zhang, 2002, for the latest discussion of model complexity and the number of free parameters). Moreover, nROUSE is a general model of perceptual identification and is readily applicable to a wide range of perceptual phenomena, including most, if not all, of the results discussed in the first section of this paper. Because of its hierarchical architecture, with feedback between levels, nROUSE is able to produce perceptual benefits, such as occurred with associative priming. The original ROUSE model supposed no such interactions between previous and current presentations, and could not handle these benefits. The perceptual hierarchy also allows nROUSE to make predictions regarding different types of features, at different levels of processing, each with their own temporal characteristics. The original ROUSE model easily handled priming data, assuming only a general pool of features, all behaving identically. However, this lack of architecture makes it difficult to predict differences for different stimulus classes using ROUSE (e.g., pattern masks vs. letter strings vs. words).

### *5.2. Other work with nROUSE*

The nROUSE model is still in a relatively early stage of development. In order to quantitatively address these short-term priming results, it was necessary to make several simplifying assumptions. Most notably, simulations were run deterministically, with non-extreme accuracy arising from a logistic function (Eq. (10)). In other simulations, we have begun the process of exploring different forms of stochastic variability. The long list of potential noise mechanisms includes, but is not limited to, selective attention, processing noise, criterial shifts, and encoding variability. The most promising possibility is the effect of stimulus differences, as realized through connection strength variability (i.e., different weight values). In support of this mechanism, Wagenmakers, Zeelenberg, and Raaijmakers (2000) observed a bias for words of higher written language frequency in forced-choice perceptual identification. nROUSE naturally explains this result, assuming high frequency words have larger connection weights. Larger connection weights cause a word to asymptote at a higher level, resulting in more rapid identification and accommodation. In any case, the most accurate noise formulation will most likely include a combination of stochastic processes and stimulus characteristics.

To inform our understanding of variability in nROUSE, we ran perceptual identification experiments requiring same/different judgment for a single test word. This allowed us to collect Receiver Operating Characteristics (ROCs), which indicate the distribution of the underlying response variable (Huber, O'Reilly, & Cousineau, in preparation). Remarkably, these experiments showed that priming causes the *z*-transformed ROC slope to exceed 1.0, indicating that the time to identify a primed target is less variable than the time to identify a primed foil. For unprimed words, the slope was less than 1.0, indicating the opposite pattern of variability. These data provide an important additional constraint in producing a reasonable formulation of the noise processes.

Finally, the shape of the reaction time (RT) distributions, for both correct and error trials, provides yet another constraint on the correct noise formulation. Because the decision rule in nROUSE is a horse race, a tight coupling is assumed between reaction time and accuracy. Specifically, persistence results in a speedup, whereas accommodation results in a slow down, for the primed alternative. Given this assumption, nROUSE makes specific qualitative predictions regarding mean RT for correct and error trials. Assuming that the time for the target to reach its peak activation largely determines RT for correct trials, nROUSE predicts that short prime durations will result in fast correct RTs whereas long prime durations will result in slow correct RTs, when the target has been primed. This same pattern is expected for error RTs when the foil has been primed, and, therefore, the relationship between correct and error RTs is opposite for priming the target as opposed to priming the foil. This pattern of RTs was observed in the single prime repetition priming experiment of Huber (in preparation). In order to explain RT distributions, Huber and Cousineau (submitted for publication) assumed target and foil peak activation times are distributed as Weibulls, and successfully fit RT distributions for the various priming conditions. Their results suggested an interaction between the target and foil, such as produced by lateral inhibition in nROUSE.

One of the strengths of nROUSE is its ability to predict activation levels at each millisecond of processing time. Temporal brain imaging techniques, such as stimulus locked Event Related Potential (ERPs), allow testing of these neural dynamics. Therefore, Huber, Curran, and O'Reilly (in preparation) performed short-term priming experiments similar to those reported in Figs. 8–10, while collecting ERPs to the target flash. By mapping the nROUSE theoretical output curves in the various experimental conditions onto the observed ERP differences, they identified signals associated with orthographic and lexical-semantic processing and observed the correlates of persistence and accommodation upon these signals. Importantly, there was a good correspondence between the dynamics of these signals and the dynamics predicted by nROUSE, both as a function of ongoing processing, and for different prime duration conditions.

### *5.3. Conclusions*

We have provided detailed fits to complex patterns of behavioral priming data using a neural-level model incorporating synapse-level accommodation mechanisms. By spanning levels of analysis in one model, we provide multiple constraints on our understanding of the dynamics of perceptual processes and neural accommodation mechanisms. These neural mechanisms allowed us to make testable predictions regarding behavioral manipulations of prime duration. In addition, our implementation of perceptual processing and accommodation is sufficiently specified to make testable predictions regarding a wide variety of other empirical manipulations.

## Note

1. The online Appendix can be found at: http://cogsci.psy.utexas.edu/supplements/.

## Acknowledgments

This research was supported by National Science Foundation (NSF) Grant KDI/LIS IBN-9873492; Office of Naval Research (ONR) Grant N00014-00-1-0246; and National Institute of Mental Health (NIMH) Grant MH63993-01. We thank Michael C. Mozer and Thomas L. Griffiths for their help with the mathematical proof that appears in the Appendix.

## References

Abbott, L. F., Varela, J. A., Sen, K., & Nelson, S. B. (1997). Synaptic depression and cortical gain control. *Science*, *275*, 220–224.

Busey, T. A., & Loftus, G. R. (1994). Sensory and cognitive components of visual information acquisition. *Psychological Review*, *101*, 446–469.

Chance, F. S., Nelson, S. B., & Abbott, L. F. (1998). Synaptic depression and the temporal response characteristics of V1 cells. *The Journal of Neuroscience*, *18*, 4785–4799.

Coltheart, M. (1980). Iconic memory and visible persistence. *Perception and Psychophysics*, *27*, 183–228.

Dixon, P., & Di Lollo, V. (1994). Beyond visible persistence: An alternative account of temporal integration and segregation in visual processing. *Cognitive Psychology*, *26*, 33–63.

Duysens, J., Orban, G. A., Cremieux, J., & Maes, H. (1985). Visual cortical correlates of visible persistence. *Vision Research*, *25*, 171–178.

Frank, M., Loughry, B., & O'Reilly, R. C. (2001). Interactions between the frontal cortex and basal ganglia in working memory: A computational model. *Cognitive, Affective, and Behavioral Neuroscience*, *1*, 137–160.

Gotts, S. J., & Plaut, D. C. (2002). The impact of synaptic depression following brain damage: A connectionist account of "access/refractory" and "degraded-store" semantic impairments. *Cognitive, Affective,* & *Behavioral Neuroscience*, *2*, 187–213.

Hogben, J. H., & Di Lollo, V. (1974). Perceptual intergration and perceptual segregation of brief visual stimuli. *Vision Research*, *14*, 1059–1069.

Huber, D. E. (in preparation). *The time course of perceptual processing for words and faces*.

Huber, D. E., & Cousineau, D. (submitted for publication). A race model of perceptual forced choice reaction time. *Proceedings of the 25th annual conference of the cognitive science society*.

Huber, D. E., Curran, T., & O'Reilly, C. (in preparation). *Using event-related potentials to test a biologically-based model of short-term priming*.

Huber, D. E., & O'Reilly, R. C. (in preparation). *Synaptic depression in short-term priming: Perceptual benefits and deficits*.

Huber, D. E., O'Reilly, C., & Cousineau, D. (in preparation). *A race model of perceptual identification 1: Reaction time distributions and receiver operating characteristics*.

Huber, D. E., Shiffrin, R. M., Lyle, K. B., & Quach, R. (2002a). Mechanisms of source confusion and discounting in short-term priming 2: Effects of prime similarity and target duration. *Journal of Experimental Psychology: Learning Memory and Cognition*, *28*, 1120–1136.

Huber, D. E., Shiffrin, R. M., Lyle, K. B., & Ruys, K. I. (2001). Perception and preference in short-term word priming. *Psychological Review*, *108*, 149–182.

Huber, D. E., Shiffrin, R. M., Quach, R., & Lyle, K. B. (2002b). Mechanisms of source confusion and discounting in short-term priming 1: Effects of prime duration and prime recognition. *Memory and Cognition*, *30*, 745–757.

Laing, C. R., & Chow, C. C. (2002). A spiking neuron model for binocular rivalry. *Journal of Computational Neuroscience*, *12*, 39–53.

Leopold, D. A., O'Toole, A. J., Vetter, T., & Blanz, V. (2001). Prototype-referenced shape encoding revealed by high-level aftereffects. *Nature*, *4*, 89–94.

Long, G. M., Toppino, T. C., & Mondin, G. W. (1992). Prime time: Fatigue and set effects in the perception of reversible figures. *Perception* & *Psychophysics*, *52*, 609–616.

Manor, Y., & Nadim, F. (2001). Synaptic depression mediates bistability in neuronal networks with recurrent inhibitory connectivity. *The Journal of Neuroscience*, *21*, 9460–9470.

Murphy, S. T., & Zajonc, R. B. (1993). Affect, cognition, and awareness: affective priming with optimal and suboptimal stimulus exposures. *Journal of Personality and Social Psychology*, *64*, 723–739.

Nelson, S. B., Varela, J. A., Sen, K., & Abbott, L. F. (1997). Functional significance of synaptic depression between cortical neurons. In J. Bowers (Ed.), *Computational neuroscience* (pp. 429–434). New York: Plenum Press.

O'Reilly, R. C. (1998). Six principles for biologically-based computational models of cortical cognition. *Trends in Cognitive Sciences*, *2*, 455–462.

O'Reilly, R. C., & Munakata, Y. (2000). *Computational explorations in cognitive neuroscience.* Cambridge, MA: The MIT Press.

Pantic, L., Torres, J. J., & Kappen, H. J. (2001). On the role of dynamical synapses in coincidence detection. *Neurocomputing*, *38/40*, 285–291.

Pantic, L., Torres, J. J., & Kappen, H. J. (2002). Associative memory with dynamic synapses. *Neural Computation*, *14*, 2903–2923.

Pitt, M. A., Myung, I., & Zhang, S. (2002). Toward a method of selecting among computational models of cognition. *Psychological Review*, *109*, 472–491.

Smith, L., & Klein, R. (1990). Evidence for semantic satiation: Repeating a category slows subsequent semantic processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition,, 16*, 852–861.

Sperling, G. (1960). The information available in brief visual presentations. *Psychological Monographs*, *74*, 1–29.

Taylor, A. L., Cottrell, G. W., & Kristan, W. B., Jr. (2002). Analysis of oscillations in a reciprocally inhibitory network with synaptic depression. *Neural Computation*, *14*, 561–581.

Taylor, T. L., & Klein, R. M. (1998). On the causes and effects of inhibition of return. *Psychonomic Bulletin* & *Review*, *5*, 625–643.

Torres, J. J., Pantic, L., & Kappen, H. J. (2002). Storage capacity of attractor neural networks with depressing synapses. *Physical Review E, 66*, 061910-5.

Tsodyks, M. V., & Markram, H. (1997). The neural code between neocortical pyramidal neurons depends on neurotransmitter release probability. *Proceedings of the National Academy of Science*, *94*, 719–723.

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

Wagenmakers, E. M., Zeelenberg, R., & Raaijmakers, J. G. W. (2000). Testing the counter model for perceptual identification: Effects of repetition priming and word frequency. *Psychonomic Bulletin* & *Review*, *7*, 662–667.

Weidemann, C. T., Huber, D. E., Shiffrin, R. M. (submitted for publication). *Spatiotemporal confusion and compensation in visual word perception*.