Cognitive Psychology 95 (2017) 79–104

Contents lists available at ScienceDirect

Cognitive Psychology

journal homepage: www.elsevier.com/locate/cogpsych

# Comparing different kinds of words and word-word relations to test an habituation model of priming

Cory A. Rieth [a], David E. Huber [b,*]

[a] *Pacific Science and Engineering Group, Inc, United States*
[b] *University of Massachusetts, Amherst, United States*

A R T I C L E  I N F O

*Article history:*
Accepted 12 April 2017
Available online 28 April 2017

*Keywords:*
Neural network modeling
Repetition priming
Associative priming
Semantic priming
Perceptual identification
Synaptic depression
Neural habituation

A B S T R A C T

Huber and O'Reilly (2003) proposed that neural habituation exists to solve a temporal parsing problem, minimizing blending between one word and the next when words are visually presented in rapid succession. They developed a neural dynamics habituation model, explaining the finding that short duration primes produce positive priming whereas long duration primes produce negative repetition priming. The model contains three layers of processing, including a visual input layer, an orthographic layer, and a lexical-semantic layer. The predicted effect of prime duration depends both on this assumed representational hierarchy and the assumption that synaptic depression underlies habituation. The current study tested these assumptions by comparing different kinds of words (e.g., words versus non-words) and different kinds of word-word relations (e.g., associative versus repetition). For each experiment, the predictions of the original model were compared to an alternative model with different representational assumptions. Experiment 1 confirmed the prediction that non-words and inverted words require longer prime durations to eliminate positive repetition priming (i.e., a slower transition from positive to negative priming). Experiment 2 confirmed the prediction that associative priming increases and then decreases with increasing prime duration, but remains positive even with long duration primes. Experiment 3 replicated the effects of repetition and associative priming using a within-subjects design and combined these effects by examining target words that were expected to repeat (e.g., viewing the target word 'BACK' after the prime phrase 'back to'). These results support the originally assumed representational hierarchy and more generally the role of habituation in temporal parsing and priming.

## 1. Introduction

Machine vision algorithms, such as the face detection software in digital cameras, take a snapshot image at a particular moment in time and then perform 'offline' calculations to determine the content of the image. However, rather than taking a discrete snapshot, biological vision involves continuous neural responses to continuously varying photoreceptor input. Furthermore, this continuous activity occurs in a staged hierarchical manner, with progressively more complex identification processes, progressing from low-level detection of dots and lines in specific retinal locations to high-level detection of

---

* Corresponding author at: Department of Psychological and Brain Sciences, 135 Hicks Way, University of Massachusetts, Amherst. Amherst, MA 01003, United States.

*E-mail address:* dehuber@psych.umass.edu (D.E. Huber).

http://dx.doi.org/10.1016/j.cogpsych.2017.04.002
0010-0285/

objects regardless of location (Hubel, 1988). Owing to the continuous nature of biological vision, there is a temporal assignment problem; in light of a constantly changing environment, identification processes need to determine whether a currently perceived feature is part of the currently viewed object versus a lingering response from a recently viewed object. To provide a concrete example, as you read each word of a sentence, you need to identify the letters of the currently viewed word without becoming confused by the letters of the most recently read word. Huber and O'Reilly (2003) proposed that neural habituation exists to solve this temporal assignment problem – by habituating to previously viewed objects, the features of prior objects are not confused with the current object.

In the lab, this temporal assignment problem is studied with Rapid Serial Visual Presentation (RSVP) paradigms (Potter & Levy, 1969). A stripped down version of RSVP presents just two stimuli in rapid succession (e.g., Duncan, Ward, & Shapiro, 1994; Hogben & Di Lollo, 1974), examining the effect of the first (the prime) on the second (the target). In the current study we used perceptual identification word priming to map out the dynamic time course of prime-target separation and we examined the manner in which this separation process differs for different kinds of letter strings, and for word pairs that are related in different ways.

For word priming, a briefly flashed prime word presented immediately before a related target typically results in better performance (i.e., faster and/or more accurate responses to the target). Reaction time distribution analyses support the hypothesis that priming benefits reflect source confusion about which letters/words appeared first versus second, with activation from the prime blending with the response to the target (de Wit & Kinoshita, 2015a; Gomez, Perea, & Ratcliff, 2013). However, the effect of this source confusion is offset, and even reversed (i.e., negative priming), if the prime is presented for a longer duration (Burt, Kipps, & Matthews, 2014; Huber, Shiffrin, Lyle, & Ruys, 2001; Huber, Shiffrin, Quach, & Lyle, 2002). This negative priming suggests an active discounting process designed to remove the unwanted effects of source confusion. Two complementary models of this discounting were developed, with each model capturing a different level of explanation (Marr, 1982): the Bayesian responding optimally with unknown sources of evidence (ROUSE) model (Huber et al., 2001) and a dynamic neural network model, nROUSE (Huber & O'Reilly, 2003). The Bayesian ROUSE model explains discounting as a near optimal perceptual decision and the nROUSE model assumes that neural habituation is the mechanism that implements discounting. Here we focus on the nROUSE model because it makes specific predictions for different prime durations. Next, we provide additional details regarding the behavioral paradigm used to measure these effects, and then we provide additional details of the nROUSE model.

### 1.1. Measuring blending and separation with immediate priming

The immediate effect of a prime word can be measured in a variety of ways, including reaction time in a lexical decision task (Meyer & Schvaneveldt, 1971; Meyer, Schvaneveldt, & Ruddy, 1974), naming latencies (Pecher, Zeelenberg, & Raaijmakers, 1998), and perceptual identification (Evett & Humphreys, 1981). Perceptual identification presents a target word very briefly followed by a mask. Participants are then asked to name the target word. The duration of the target is set separately for each individual to place accuracy at a threshold level (e.g., 75% correct identification). This task has some advantages over lexical decision for studying the perceptual separation of prime and target because it requires overt identification of the target, which directly assesses automatic processes (Pecher, Zeelenberg, & Raaijmakers, 2002). In contrast, lexical decisions are influenced by a sense of familiarity (Wagenmakers, Zeelenberg, Steyvers, Shiffrin, & Raaijmakers, 2004) or retrospective semantic matching (de Wit & Kinoshita, 2015b). Rather than asking participants to name the briefly flashed target word, control over the decision process is gained through a forced-choice decision, with accuracy providing the key measure of interest (Ratcliff & McKoon, 1997; Ratcliff, McKoon, & Verwoerd, 1989), although reaction time produces analogous results (Huber & Cousineau, 2004; Potter, Donkin, & Huber, submitted for publication).

Huber et al. (2001) examined immediate priming in a perceptual identification paradigm with forced-choice testing to determine if priming produces both costs and benefits. Fig. 1 presents a modified version of this procedure as used by Huber (2008) and the current experiments. After an initial fixation cross, a prime word was presented. On each trial, the duration of the prime was randomly chosen from one of five durations, ranging from a subliminal presentation (17 ms) to an excessively long presentation (2000 ms). The prime word was shown in a doubled up fashion (i.e., two copies of the prime word, with one on top of the other). This was done to provide some visual difference between prime and target for the case of repetition priming, although similar results are found when using lower case primes followed by upper case targets (see Experiments 2 and 3 of Huber et al., 2001). Participants were informed that the prime was just as likely to indicate the wrong answer as the right answer and this assertion was reinforced through trial-by-trial accuracy feedback. Thus, it was made clear that the prime was non-diagnostic (i.e., could not be used to help guess the correct answer). Immediately following the prime, a target word appeared for a duration set individually so that performance was 75% on average (i.e., at the perceptual threshold). After the target presentation, a pattern mask appeared, followed by a choice between the correct target word and an incorrect foil word. This response was non-speeded and accuracy was the measure of interest.

Using this procedure, Huber et al. (2001) observed both costs and benefits of priming as measured with identification accuracy. Furthermore, increases in prime duration caused a reversal in the pattern of costs and benefit. With short duration primes, repetition priming of the target (i.e., the 'target primed' condition) produced better performance as compared to a condition with unrelated choices (i.e., the baseline 'neither primed' condition). However, with repetition priming of the foil (i.e., the 'foil primed' condition), performance was worse than the baseline neither primed condition. This pattern describes 'positive priming' because priming either the target or the foil increased the likelihood that the participant would choose the

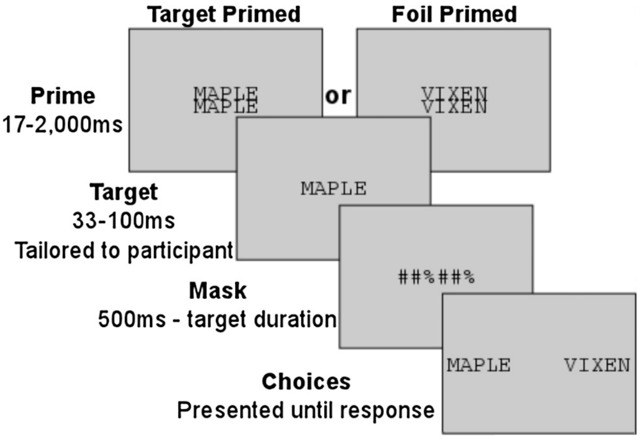


**Fig. 1.** Basic structure of a trial. After an initial fixation, a prime appeared for one of five durations, ranging from 17 to 2000 ms. The relationship between the prime and the primed choice word varied between experiments. In Experiment 1, the prime was identical to one of the choice words, whereas in Experiment 2, the prime was a semantic associate of the primed choice word. In Experiment 3, the prime was either a repetition of a primed choice word or a different word likely to precede the primed choice word according to a written corpus analysis. After the prime, the target was displayed for a duration (between 33 and 100 ms) set separately for each individual to place accuracy at a 75% threshold. The target was immediately followed by a mask, presented for a duration such that the choice words appeared 500 ms after the onset of the target. Finally, two choice words were presented, left and right of where the target appeared, one identical to the target, and another an incorrect foil. Experiments 1 and 2 used a doubled upper-case prime word (as shown) to provide some visual difference between prime and target in the case of repetition priming. In contrast, Experiment 3 presented a single prime in lower-case letters in same midline position as the upper case target.

primed alternative. With long duration primes, this entire pattern flip-flopped, with target primed worse than neither primed while foil primed was better than neither primed. This describes ‘negative priming’ because the participant was less likely to choose the primed alternative. At first glance, this might appear to be a strategic decision bias. However, it is not clear why the decision bias should reverse with prime duration. Furthermore, there is evidence that these effects occur automatically (Huber, Shiffrin, Lyle, & Quach, 2002; Weidemann, Huber, & Shiffrin, 2008), reflecting early perceptual responses (Huber, Tian, Curran, O’Reilly, & Woroch, 2008).

### *1.2. The nROUSE model*

The change from positive to negative priming was explained by the Bayesian ROUSE model as a consequence of under- versus over-discounting of the prime in the attempt to correct for perceptual source confusion (Huber et al., 2001). However, the Bayesian ROUSE model does not explain why there is a gradual transition from under- to over-discounting with increasing prime duration, requiring a separate discounting parameter value at each prime duration. To explain the time course, Huber and O’Reilly (2003) developed a neural dynamics model, termed nROUSE. The nROUSE model explains temporal separation in general, with a particular application to word priming. To date, nROUSE has been applied to word priming (Huber, 2008) and word priming electrophysiology (Huber, Tian, Curran, O’Reilly, & Woroch, 2008), face priming (Rieth & Huber, 2010), immediate change detection (Davelaar, Tian, Weidemann, & Huber, 2011), recognition priming (Huber, Clark, Curran, & Winkielman, 2008), semantic satiation (Tian & Huber, 2010), semantic satiation electrophysiology (Tian & Huber, 2013), continuous flash suppression (Huber, 2015), the priming of affectively valenced pictures (Irwin, Huber, & Winkielman, 2010), and the attentional blink (Rusconi & Huber, submitted for publication).

There are three critical components of the model as applied to this word priming paradigm:

(1) DYNAMICS: habituation in the form of activity-dependent synaptic depression.
(2) REPRESENTATION: a perceptual cascade structure capturing different levels of perception (e.g., visual, orthographic, and lexical-semantic in the case of words; see Fig. 2).
(3) DECISION: a race process based on lexical-semantic identification latencies determines choice behavior.

The mathematical details of the nROUSE model as currently applied are reported in Appendix A.

The nROUSE model explains the transition from positive to negative priming in the following manner. Input is provided to the visual layer of the model at each millisecond of the trial sequence as dictated by the types of stimuli and durations appropriate to the simulated condition. Activation arising from short duration primes is substantial because there is insufficient time for habituation, and this residual activation causes faster identification of a primed choice word (i.e., the primed choice word has a head start in the decision process). Thus, accuracy is increased for target priming but decreased for foil priming, both as compared to a condition in which the prime is unrelated to either choice. Following a longer duration prime, residual activation from the prime is reduced, owing to habituation. In the case of associative priming, this reduces the magnitude of

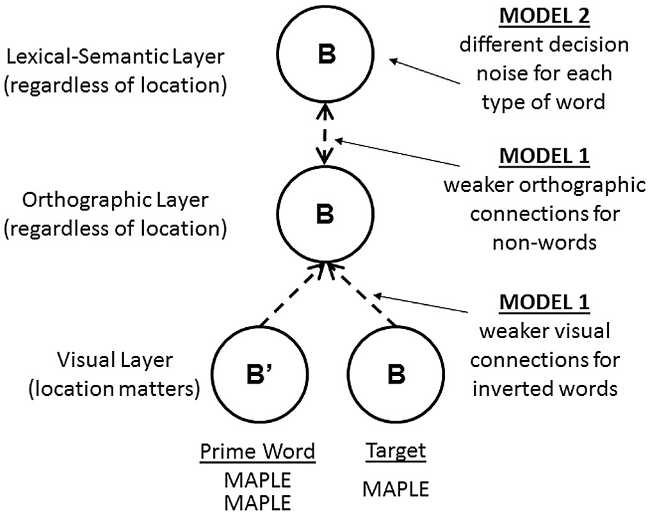


**Fig. 2.** The assumed representations of the 3-layer nROUSE model, with artificial neurons capturing the dynamic activity of early visual processes in response to letter strings, which in turn drive orthographic and then lexical-semantic representations. The figure shows the connectivity for repetitions (a target word labeled B following a repetition prime presented in a different visual location, B′). The visual layer is assumed to be retinotopic. Thus, the same letter string presented in different screen locations entails separate visual layer neurons. In contrast, orthographic and lexical-semantic representations are assumed to have large receptive fields such that the same neurons are activated regardless of where the word appears. For this reason, lingering orthographic and lexical-semantic activity from the prime word blends with activity from the target or choice words. However, habituation reduces this blending, and produces negative priming in the case of repetition priming. As compared to known words, non-words should have a weakened connection between the non-word's orthography and any lexical-semantic representation evoked by the non-word. In contrast, inverted words should have a weakened connection between the inverted word's visual response and the associated orthographic response. Feedforward connection strengths are by default set to 1.0, with feedback from the lexical-semantic layer set to 0.25 times the value of the forward connection. The weakened feedforward connections should take on values less than 1.0, as indicated by dashed lines connecting a pair of artificial neurons. Thus, according to MODEL 1, performance should be worse for these types of letter strings and, concomitantly, the transition from positive to negative priming should be slower owing to weaker connections. This model was contrasted with MODEL 2, in which the different types of letter strings were assumed to have the same connectivity as known words, but differ in decision noise.

positive priming. In the case of repetition priming, which reuses the same (now habituated) feed-forward connection between the repeated orthography and its meaning, lingering habituation slows identification latency, causing decreased accuracy for target priming as well as increased accuracy for foil priming. In short, because the orthography of a repeated word is slow to re-activate at the time of the forced-choice decision, nROUSE produces negative priming.

### *1.3. The predicted relationship between connection strength and habituation*

In the nROUSE model, habituation through synaptic depression (Tsodyks & Markram, 1997) is achieved by assigning each simulated neuron two dynamic variables (see Appendix A). The first is the voltage potential, capturing the effect of excitatory and inhibitory input to the cell. This variable is akin to activation in a traditional connectionist model, specifying the probability that the cell will undergo an action potential (i.e., average firing rate). The second dynamic variable specifies the effect of each action potential. More specifically, if no neurotransmitter is available for release, then the action potential will have no influence on receiving cells. This second variable indicates the current level of available neurotransmitter and is multiplied by the firing rate to determine the output of the cell. If the cell has recently been active, its supply of neurotransmitter will be depleted, and it will be unable to communicate with receiving cells (i.e., it may be attempting to shout, but has temporarily lost its voice).

Habituation through synaptic depression predicts an interaction between connection strength and the time course of habituation. If a sending cell is strongly connected to a receiving cell, then it will more strongly excite the receiving cell initially, but it will also more rapidly deplete its reserve of neurotransmitter. By analogy, the changed dynamics of habituation are like emptying a glass of water versus emptying a water bottle; the narrow neck of the bottle (i.e., a weak connection) provides a weak initial flow of water (i.e., an initially poor perceptual response), but, as a result, the water bottle depletes its reservoir of water more slowly and can provide a flow of water for a longer duration (i.e., slower habituation).

### *1.4. The current study*

Although the main focus of the current study was the predicted relationship between connection strength and habituation, a test of this prediction necessitated an assessment of the model's representational assumptions. Aside from habituation through synaptic depression, development of the nROUSE model required auxiliary assumptions regarding representation (a 3-layer cascade) and perceptual decision making (a race process). A separate study addressed the decision making assumptions of the model (Potter et al., submitted for publication), examining reaction time distributions and applying a sequential sampling race model to determine separate drift rates for target and foil choice words in each priming condition. Complementing that study, the current experiments addressed the representational assumptions of the nROUSE model. The assumed three-layer hierarchy produces a perceptual cascade in which lexical-semantic activation occurs as soon as orthographic representations become active, similar to the interactive-activation model (McClelland & Rumelhart, 1981).

In the current study, we manipulated the stimuli in an attempt to compare situations with stronger connectivity (e.g., valid words) to situations with weaker connectivity (e.g., non-words). If the assumed representation is valid, then the priming dynamics are predicted to reveal faster habituation for stronger connections. Because a repeated word repeats both the orthography and semantics, prior applications of nROUSE to repetition priming did not manipulate connection strength or otherwise test the assumed representation. Thus, we compared different kinds of words and different kinds of word-word relations, with these manipulations predicted to change the dynamics of priming in specific ways under the assumed representation. However, using the observed priming results to test the representational assumptions is an example of the logical fallacy known as affirming the consequent (e.g., if representational assumptions true, then faster habituation predicted; faster habituation observed, thus representational assumptions are true). The obvious remedy is to consider other accounts of the observed priming results (i.e., other representational assumptions). Thus, for each experiment, we engaged in model comparison, testing the adequacy of the original hierarchal representation (MODEL 1) to plausible alternative representations (MODEL 2).

Experiment 1 compared repetition priming for different kinds of stimuli assumed to differ in connection strength. Because this experiment included a large number of conditions, the results were used to determine model parameter values appropriate for the display characteristics and font used in this line of experiments. Except for parameters related to the properties of the stimuli that varied between experiments, the only parameter allowed to vary across experiments was the decision noise parameter, which was allowed different values considering that each experiment used slightly different procedures for determining the threshold target duration (different thresholding procedures may have produced overall higher or lower accuracy). Critically, the decision noise parameter does not affect the dynamics of the model; instead, it merely rescales the results of the perceptual race process into the accuracy scale in a mirror symmetric fashion about the 50% chance level (i.e., if the target wins, the degree by which the target wins specifies an accuracy between 50% and 100%, and if the foil wins, the degree by which the foil wins specifies an accuracy between 50% and 0%). In determining the model parameters, it was fit to the results after averaging over subjects, but we note that individual differences were small (see Appendix B)[1] and these experiments used a large number of subjects; thus, the best-fitting parameters from one experiment should provide an adequate account of the other experiments.

This fixing of parameters across experiments is crucial because all else being equal, the representational assumptions for associative priming (as revealed in Experiment 2) predicted less habituation as compared to repetition priming (as revealed in Experiment 1). This prediction is made because the target's orthography is not habituated by an associatively related prime. Thus, by fixing the parameters across experiments, the model provided an internally consistent account of associative priming in relation to repetition priming, testing whether the difference between the types of priming can be captured solely as a function of a single free parameter specifying the semantic similarity between a prime and an associatively related target (whereas semantic similarity is 100% for repetitions). Because it was conceivable (but unlikely) that the observed differences between experiments reflected some sort of strategy difference when confronted only with repetition priming or only with associative priming, Experiment 3 tested both repetition priming and semantic priming using a within-subjects design. In addition, Experiment 3 examined the combination of these factors with words that were expected to repeat based on natural language co-occurrence statistics.

## 2. Experiment 1

Experiment 1 compared repetition priming across 4 different kinds of letter strings (words, pronounceable non-words, non-pronounceable non-words, and inverted words) that were expected to entail different connection strengths in the assumed hierarchy. For each of these different types of letter strings, different prime durations were tested to map out the time course of the transition from positive to negative priming. Non-words can be read (i.e., they have an orthographic representation) but they evoke lexical-semantic representations only by orthographic/phonological similarity to known words. Thus, as compared to words, non-words should have a weakened connection between orthographic and lexical-semantic representations. Inverted words can be read as well, albeit slowly, but in this case the connection between orthographic and lexical-semantic representations should be intact. Instead, the difficulty with inverted words stems from the

---

[1] nROUSE can be fit separately to each subject with sufficient data per subject (Potter et al., submitted for publication). However, the current study only collected 10 observations for each condition for each subject. Providing evidence of the minimal role of individual differences, we report in Appendix B the subject-level results of an experiment that examined a subset of the Experiment 1 word conditions with 100 trials per condition for each subject. As seen in Appendix B, every subject exhibited positive priming for short duration primes and the magnitude of priming was reduced for longer duration primes for every subject. Subject differences in this paradigm are minimal in large part because separate target durations are set for each subject to place performance at the same part of the accuracy scale.

mapping between the visual input and orthography. With these assumptions, nROUSE predicts that the transition from positive to negative priming should be slower non-words and inverted words as a function of increasing prime duration.

### 2.1. Method

#### 2.1.1. Participants

A total of 63 undergraduate psychology students participated in this experiment for course credit.

#### 2.1.2. Equipment

All three experiments were conducted on LCD monitors with a 60 Hz refresh rate. Presentation times were synchronized with the display refresh rate. All items were displayed on a gray background and the display resolution was set to 640 by 480 pixels.

#### 2.1.3. Stimuli

There were four stimulus types in the experiment: words, pronounceable non-words, non-pronounceable non-words, and inverted words. Stimuli were created starting with a 746 word list of syllable-marked two-syllable five or six letter words. Words were hand selected to eliminate low frequency items. Importantly, non-words were created from components of the actual words so that each letter appeared in each position the same number of times for each stimulus type over the course of the experiment. Thus, letter frequency/position was controlled. This is a critical control, as letter representations are sensitive to letter position (New & Grainger, 2011). Non-word stimuli were created by combining pairs of the syllables appearing in the actual words to create non-words and non-pronounceable non-words. Starting with a pair of words, two potential pronounceable non-words were created by switching syllables between the two words. For example, from the words ‘rustic’ and ‘radish’ the pronounceable non-words ‘rudish’ and ‘rastic’ were created. Potential non-pronounceable non-words were created by selecting alternating letters from the word pair. In the previous example, the corresponding non-pronounceable non-words would be ‘rudtsc’ and ‘rasiih.’ Non-word status was confirmed with a 410,000 word English dictionary (containing words as well as abbreviations, alternate spellings and common names). In addition to preserving syllables for the pronounceable non-words, it was required that all trigrams (3 letter combinations) within a pronounceable non-word exist as part of at least one word in the dictionary. In contrast, for non-pronounceable non-words it was required that at least one trigram not exist as part of any word in the dictionary. All non-words used in the experiment were additionally checked by hand to ensure they were suitable. Stimuli for the inverted word condition were simply the word stimuli rotated 180 degrees. This procedure resulted in 434 unique stimuli for each condition. The stimuli were randomly assigned without replacement to prime durations and priming type for each participant. Furthermore, the targets and foils for each trial were matched to have the same number of letters on a particular trial so that target length was not a reliable distinguishing feature. All stimuli were presented in Courier. Table 1 presents example stimuli of each type for each stimuli type and priming condition.

#### 2.1.4. Procedure

Fig. 1 presents an overview of the trial paradigm. The task was to identify the briefly flashed central target word as assessed by a forced-choice decision. Primes were presented immediately prior to the target in the center of the screen for durations ranging from 17 ms (i.e., subliminal priming) to 2000 ms (i.e., excessive priming). In this manner, the prime word also served as a forward mask of the target word. Responses were self-paced and accuracy feedback was provided on every trial. Participants were informed that the prime would be identical to the target as often as it would be identical to the incorrect answer such that there could be no effective decision strategy to choose or not choose the primed alternative.

Each trial began with a fixation cross presented for a duration equal to 2500 ms minus the prime duration for that trial. This maintained a constant duration between initiation of the trial and presentation of the target such that participants could know when to expect the target, equating attentional alertness to the briefly flashed target regardless of prime duration. Next, two identical copies of the prime were simultaneously presented, one directly above the other. Immediately after the prime, a target word was briefly presented in the center of the screen followed by a pattern mask consisting of random arrangements of the symbols ‘#’ and ‘%’. The mask was presented for a duration equal to 500 ms minus the target duration. This maintained a constant duration between onset of the target and onset of the choice alternatives. The reason to do so is to avoid confounding stimulus onset asynchrony between target onset and choice word onset across individuals who differ in the target duration necessary for threshold performance. Following the mask, the target and foil words were presented on the left and right sides of the screen, separated by approximately 18 degrees of visual angle. Responses were collected through a button box, with the far left or right buttons corresponding to the selection of the left or right option on the screen. The two choices were presented until a response was made, after which accuracy feedback was presented for 1500 ms. The choice position of the target word was randomly counterbalanced between the left and right sides.

After instruction, participants completed a block of 16 practice trials. These 16 trials included four trials of each of the four stimulus types (word, pronounceable non-word, non-pronounceable non-word, and inverted word), in random order. For all 16 trials, the prime was unrelated to the choices. For these practice trials, the prime duration was fixed at 500 ms and the duration of the target was gradually reduced from 300 ms to 100 ms across trials. If participants failed to achieve accuracy

**Table 1**
Conditions and example stimuli for Experiment 1.

| Stimuli type | Primed choice | Example prime | Example target | Example foil |
|---|---|---|---|---|
| Words | Target | MAPLE | MAPLE | VIXEN |
| Words | Foil | VIXEN | MAPLE | VIXEN |
| Pronounceable non-words | Target | PAYEN | PAYEN | SKIKA |
| Pronounceable non-words | Foil | SKIKA | PAYEN | SKIKA |
| Non-pronounceable non-words | Target | BNSEN | BNSEN | MATLL |
| Non-pronounceable non-words | Foil | MATLL | BNSEN | MATLL |
| Inverted words | Target | DOGMA | DOGMA | TAFFY |
| Inverted words | Foil | TAFFY | DOGMA | TAFFY |

greater than 75%, they were shown the instructions again, followed by another block of 16 practice trials. This procedure of repeating the instructions followed by 16 practice trials repeated until they achieved over 75% accuracy.

To accommodate individual differences, the target duration was determined separately for each participant to obtain 75% accuracy. This target duration was determined based on a block of 80 trials. Sixteen conditions were tested five times each, in random order, during this block of trials. These 16 conditions represented all combinations of 4 different target durations (33, 50, 67, and 83 ms) crossed with the four stimulus types (word, pronounceable non-word, non-pronounceable non-word, and inverted word). As in the practice trials, the prime was always unrelated to both choices. Accuracy during the 80 trials was collapsed across stimulus type, resulting in four target duration conditions. Based on these four conditions, participants were assigned the lowest target duration at which they answered 75% of the trials correct. If 75% identification was not achieved with the 83 ms target, a target duration of 100 ms was assigned. The remainder of the experiment used the target duration determined by this procedure.

After determination of an appropriate target duration, the main portion of the experiment ensued. Three variables were manipulated within subject: five levels of prime duration (17, 50, 150, 400 or 2000 ms), primed choice with two levels (target primed or foil primed), and stimulus type with four levels (words, pronounceable non-words, non-pronounceable non-words, or inverted words). In the target and foil primed conditions, the prime was identical to either the target or foil, respectively. For a trial of a given stimulus type, all presentations during the trial were of that type (e.g., an inverted prime followed by an inverted target followed by inverted choice words). Participants were given breaks between blocks. In total, the experiment lasted approximately 35–45 min.

There were two between-subjects variants of the experiment. In the first (N = 29 participants), stimulus type was randomly counterbalanced within each block, and there were five blocks of 80 trials each (each block consisted of 2 trials of each of the 40 conditions). In the second variant (N = 34 participants), stimulus type was consistent within each block, and there were four blocks of 100 trials each (i.e., each block consisted of 10 trials for each of the 10 conditions of a single stimulus type). For this variant, block order was randomly determined for each participant. Regardless of whether stimulus type was blocked or random, there were 10 trials for each condition in total across the experimental session. Statistical tests revealed that the transition from positive to negative priming was slightly faster with the blocked presentation of stimulus type. However, blocked versus random did not interact with any of the other variables, or produce any higher order interactions, and so the results were collapsed over the blocking manipulation.

## 2.2. Results

### 2.2.1. Performance differences

Statistical analysis of performance during the threshold portion of the experiment revealed main effects of stimulus type ($F(3,186) = 27.72$, $p < 0.001$, $\eta^2 = 0.065$) and target duration ($F(3,186) = 102.00$, $p < 0.001$, $\eta^2 = 0.183$). Averaged over target durations, accuracy was higher for word stimuli than all other stimuli types (all $t(251) \geq 2.43$, $p_{holm} \leq 0.031$; all reported $p$-values from $t$-tests are Holm adjusted for multiple comparisons). Performance was higher for pronounceable non-words than either non-pronounceable non-words or inverted words ($t(251) \geq 2.03$, $p_{holm} \leq 0.043$). Lastly, non-pronounceable non-words were more accurately identified than inverted words ($t(251) = 5.05$, $p_{holm} < 0.001$).

### 2.2.2. Priming effects

We define a priming effect as the difference between the target primed and foil primed conditions (i.e., an effect of primed choice), which could be positive, negative, or absent. Showing results from the main portion of the experiment, Fig. 3 presents average accuracy for all combinations of prime duration, primed choice, and stimulus type, as seen in the left-hand set of graphs, as well as the priming effects, as seen in the right-hand set of graphs. In a repeated measures ANOVA comparing the factors of primed choice (i.e., priming effect), prime duration, and stimulus condition on proportion correct, there was a significant three way interaction ($F(12,744) = 5.08$, $p < 0.001$, $\eta^2 = 0.012$). All lower order interactions and main effects were also significant. Exploring the three way interaction, all stimulus types produced a positive priming effect for prime durations of 17, 50, and 150 ms (all $t(62) \geq 4.38$, $p_{holm} < 0.001$). For 400 ms primes, there was no significant priming effect for words, pronounceable non-words, and non-pronounceable non-words (all $|t(62)| \leq 1.96$, $p_{holm} > 0.27$), but there was a

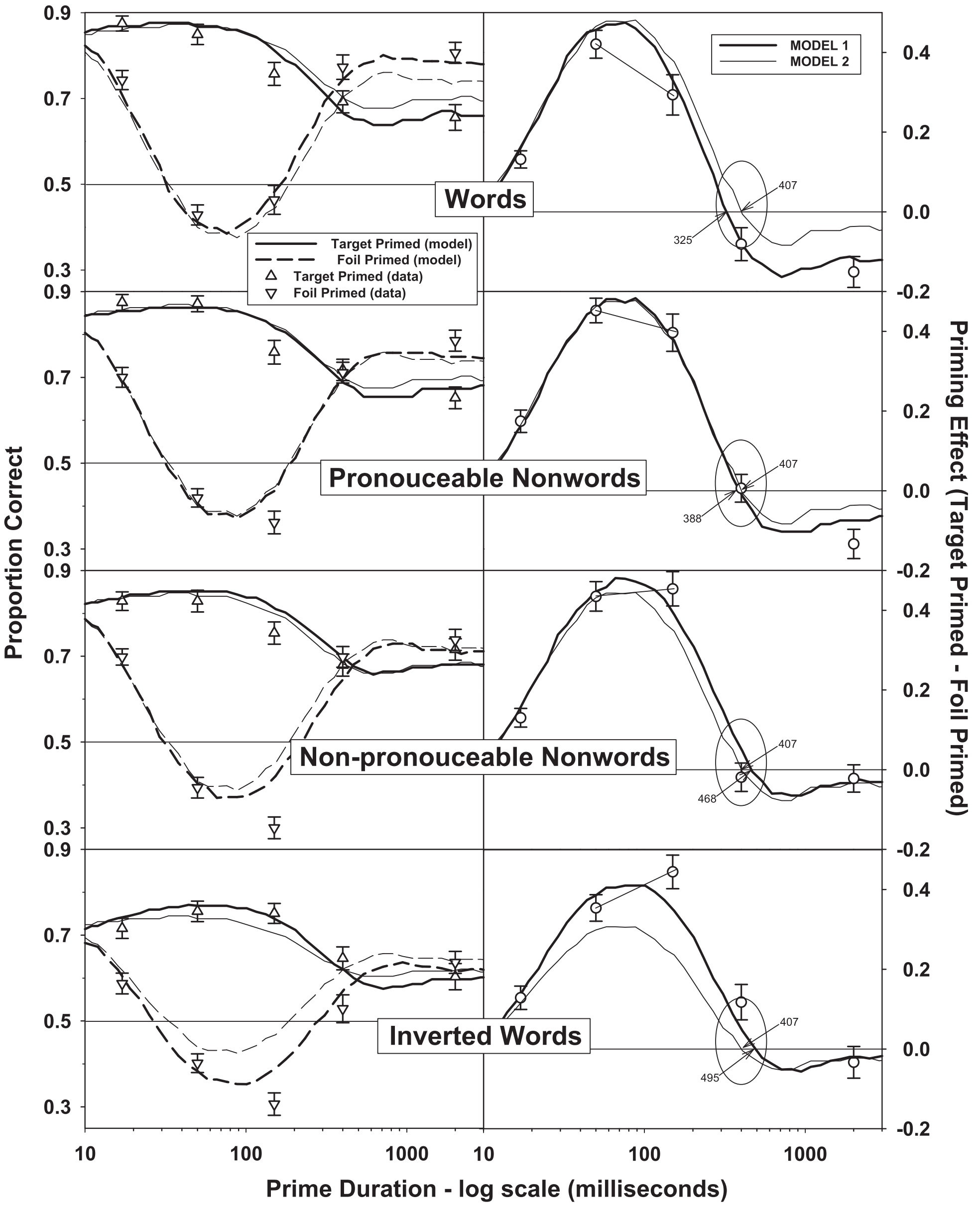


**Fig. 3.** Experiment 1 results (symbols) and model results (lines). The left set of graphs shows average accuracy in each condition and the right set of graphs shows priming effects by taking the difference between the target primed and foil primed conditions. Error bars are +/−1 standard error of the mean, or standard error of the mean difference. The 40 conditions of this experiment were fit using 7 free parameters for each model, with MODEL 1 producing a $G^2$ error of 155.5 versus 300.7 for MODEL 2. For the right-hand set of graphs, the lines connecting the 50 ms and 150 ms priming effects of the observed data highlight changes across the four stimulus types in terms of the prime duration that produced the largest priming effect. The ovals at 400 ms, which are centered at a priming effect of 0, highlight the changes in the direction of priming when comparing different kinds of stimuli. For each model, the prime duration producing a transition from positive to negative priming is indicated. The observed changes in the priming dynamics were well explained by MODEL 1, whereas MODEL 2 rescales the magnitude of the priming effect for each type of stimuli, but assumes the same priming dynamics.

marginally significant positive priming effect ($t(62) = 2.65$, $p_{holm} = 0.061$) for inverted words. For 2000 ms primes, there was a significant negative priming effect for words ($t(62) = -3.90$, $p_{holm} = 0.002$) and pronounceable non-words ($t(62) = -3.63$, $p_{holm} = 0.004$), but no significant effect for non- pronounceable non-words ($t(62) = -0.65$, $p_{holm} = 1$), or inverted words ($t(62) = -0.84$, $p_{holm} = 1$).

Next, we examined priming effects as prime duration increased. We found increased priming for all stimuli types as prime duration increased from 17 ms to 50 ms (all $t(62) \leq -5.90$, $p_{holm} < 0.001$). Word priming decreased from the 50 ms to the 150 ms prime durations ($t(62) = 2.83$, $p_{holm} = 0.037$) but there were no priming effect changes for the other three types of priming (all $|t(62)| \leq 1.23$, $p_{holm} \geq 0.662$). From 150 ms to 400 ms, the priming effect decreased significantly for all stimuli types, (all $t(62) \geq 6.93$, $p_{holm} < 0.001$). There was no significant change from 400 ms to 2000 ms for words ($t(62) = 2.45$, $p_{holm} = 0.085$) or non-pronounceable non-words ($t(62) = 0.10$, $p_{holm} = 1$), whereas there was a significant decrease for pronounceable non-words ($t(62) = 3.39$, $p_{holm} = 0.010$) and inverted words ($t(62) = 3.36$, $p_{holm} = 0.010$).

In summary, the rank ordering in terms of performance was words > pronounceable non-words > non-pronounceable non-words > inverted words. Furthermore, this same ordering captured the transition from positive to negative priming (i.e., words had the fastest transition and inverted words the slowest).

### *2.2.3. Model comparison*

The reported ANOVA and post hoc comparisons provide a crude assessment of whether the observed data conformed to predictions. A better test is to engage in model comparison to determine the extent to which the data are better captured by a model that contains the predicted changes in the priming dynamics as compared to one that fixes the priming dynamics. As seen in Fig. 2, if the second and third layers of the model correspond to orthographic and lexical-semantic representations, then non-words should entail weaker connections from layer 2 to layer 3 and inverted words should entail weaker connections from layer 1 to layer 2. This MODEL 1 used the same parameters for all conditions except for a single changed layer 2–3 connectivity parameter for each kind of non-word and a single changed layer 1–2 parameter for inverted words. As seen in Fig. 3, these changed connectivity parameters produced both lower performance (the left-hand graphs) and slower transitions from positive to negative priming with increasing prime duration (the right-hand graphs). The results are plotted using a log scale for prime duration considering that the experimental design used prime durations chosen with approximately equal spacing on a log scale (with a linear scale the short prime duration conditions would be sandwiched on top of each other). However, this scale makes it difficult to see the extent to which the model predicts different positive-to-negative priming transitions for each class of stimuli. Thus, the figure also includes annotations of the predicted crossover points, showing that for MODEL 1, words crossover to negative priming at a prime duration of 325 ms whereas inverted words crossover to negative priming at a prime duration of 485 ms (i.e., a slowdown of 160 ms).

MODEL 1 was compared with MODEL 2, which contained the same number of free parameters. However, in MODEL 2, all of the connections were set at the default value of 1.0 and lowered accuracy in the non-word and inverted word conditions was captured by allowing the decision noise parameter to take on a different value for each type of stimulus. Critically, although decision noise reduces accuracy, it leaves the dynamic transition from positive to negative priming unchanged; MODEL 2 rescales accuracy in a mirror symmetric fashion about the chance level of 50%. Thus, a goodness of fit comparison between MODEL 1 and MODEL 2 provided a statistical test of whether the dynamic time course of priming differed across the different types of stimuli. As seen in the figure, the primed duration producing a crossover to negative priming for MODEL 2 was 407 ms for all classes of stimuli.

Average accuracy for the 40 conditions of Experiment 1 were fit by both MODEL 1 and MODEL 2 using 7 free parameters to minimize the $G^2$ error, which consequently maximizes the likelihood of the data (Riefer & Batchelder, 1988). The best-fitting $G^2$ error was 155.5 for MODEL 1 whereas it was 300.7 for MODEL 2 (see Appendix A for additional details of fitting the models). As seen in the left-hand set of graphs, both models captured accuracy differences in general, with inverted words revealing the worst performance and known words revealing the best performance. However, as seen in the right-hand set of the graphs, MODEL 2 failed to capture the observed differences in priming dynamics that occurred when comparing the different kinds of stimuli. The lines connecting the observed priming effects following a 50 ms prime versus a 150 ms prime highlight the successes of MODEL 1 and the failures of MODEL 2. As seen with these lines, when increasing the prime duration to 150 ms, words produced a large decrease in priming, pronounceable non-words produces a modest decrease in priming, non-pronounceable non-words produced a modest increase in priming and inverted words produced a large increase in priming. This same ordering is apparent as highlighted by the ovals at the 400 ms prime duration: known words produced negative priming, non-words produced absent priming, and inverted words produced positive priming. MODEL 1, with its change in the connectivity parameters, captured these effects whereas MODEL 2 only rescaled the magnitude of priming at each prime duration but could not capture changes in the time course of the transition from positive to negative priming.

## *2.3. Discussion*

Experiment 1 confirmed key predictions of the nROUSE habituation model by examining repetition priming for different types of stimuli designed to capture different strengths of visual, orthographic, and lexical-semantic representation. Sensibly, performance was better for words than non-words, and, as predicted, words exhibited a faster transition from positive to negative priming than non-words. The model comparison results supported these conclusions and a model assuming changed connectivity (MODEL 1: see Figs. 2 and 3) was able to capture the changes in the dynamics of priming when comparing the different kinds of stimuli whereas MODEL 2 could not, despite capturing the on-average accuracy differences across the stimuli and having the same number of free parameters as MODEL 1. Thus, the observed data, and the model comparison results support the assumption that layer 2 of the nROUSE model represents orthographic representations and layer 3 of the nROUSE model represents lexical-semantic representations.

These results are consistent with other results in the literature. For instance, the inverted word results are similar to the finding that upright faces reveal a more rapid transition to negative repetition priming as a function of increasing prime duration as compared to inverted faces (Rieth & Huber, 2010). However, upright faces are often considered a ‘special’ stimulus (Yue, Tjan, & Biederman, 2006), and may be expected to have unique processing characteristics relative to inverted faces (Yin, 1969) beyond how well-learned they are as stimuli. That these inversion priming effects also occurred with inverted words suggests that the face inversion priming results did not occur because of a qualitative change in the nature of the stimuli (at least as it relates to this priming paradigm, inverted faces are not special).

The non-word results are similar to the results of Humphreys, Besner, and Quinlan (1988). Similar to the prime duration effects reported here, that study found that masked primes produced positive priming and unmasked primes produced negative priming. Furthermore, they found that the magnitude of negative priming for unmasked primes was larger for high frequency words as compared to low frequency words (pronounceable non-words can be thought of as very low frequency words). However, because their study only examined two prime durations, it is difficult to distinguish between overall changes in the magnitude of priming versus changes in the rate at which positive priming transitioned to negative priming with increasing prime duration.

In summary, Experiment 1 confirmed the prediction that well known stimuli should produce better performance and a faster transition to negative repetition priming whereas less well known stimuli produce worse performance and a slower transition to negative repetition priming. Furthermore, these results were quantitatively captured by the nROUSE model by changing representational strength in a manner consistent with the types of stimuli in each condition.

## 3. Experiment 2

The experimental trials of Experiment 2 were identical to the word trials of Experiment 1 except for the use of associative primes rather than repetition primes. Because associative/semantic priming does not involve any shared orthographic representations, the model predicted a rise and fall of associative priming as a function of increasing prime duration, but fail to crossover to the negative priming observed with repetition priming. Huber (2008) examined the time course of associative priming with this paradigm, but that study only tested the both primed condition (e.g., prime word GRAVY and choice words SAUCE versus TRAIN) rather than the target primed and foil primed conditions necessary for mapping out the rise and fall of priming.

MODEL 1 in Fig. 4 shows the manner in which Huber (2008) addressed associative priming. For instance, consider the prime word RAIL followed by the target word ROAD. The prime and target are assumed to have distinct orthographic representation (A versus B at layer 2), but a shared semantic representation (Fischler, 1977; Lucas, 2000). This shared meaning is indicated by the intersection of A and B shown in the dashed circle of Fig. 4 (e.g., the meaning of ‘railroad’). In addition, each word has some semantic representation that is distinct from the shared meaning (e.g., the various meanings of ROAD other than ‘railroad’). With habituation, the activation of the shared meaning is lessened, which reduces the magnitude of priming (the difference between the target primed and foil primed conditions should be reduced). However, because there is no connection evoked by both the prime and target, associative priming should not become negative. This is because true deficits only occur in the nROUSE model when attempting to reuse an already depleted connection between representations. For instance, seeing the word RAIL may initially make you think about railroads and with extended viewing of RAIL you eventually think less about railroads but that meaning is not inhibited. Instead, this only reduces the positive priming effect for the target word ROAD rather than making it difficult to read ROAD. In contrast, if you had seen ROAD for an extended period of time as a repetition prime, then the connection between the orthography of ROAD and the meaning of road would become depleted, and it would be difficult to access the meaning of ROAD when it appeared in a new location.

If associative priming reflects shared meaning, then that shared meaning should be accessible from either prime or target. Thus, it should be possible to obtain priming effects even if the order of events is reversed, switching the prime and target words. Word association strength is experimentally determined by presenting participants with a cue word and asking them to produce the first word that comes to mind (Nelson, McEvoy, & Schreiber, 1998). Many associated word pairs are bi-directional such that each is likely to result in the production of the other (e.g., ‘light’ and ‘dark’). However, some word pairs are unidirectional, with an association only in one direction (e.g., in response to the cue ‘saber’, participants will produce ‘tooth’, but when cued with ‘tooth’ they never respond ‘saber’). By comparing forward (’saber’ priming ‘tooth’) versus backward (’tooth’ priming ‘saber’) priming, unidirectional word pairs were used to examine whether priming is directional in this perceptual identification paradigm. According to MODEL 1, both ‘saber’ and ‘tooth’ should elicit thoughts of saber-toothed tigers and differences between forward versus backward priming should be minimal both in priming magnitude and in the time course of priming with increasing prime duration.

### *3.1. Method*

### *3.1.1. Participants*

A total of 73 undergraduate psychology students participated in this experiment for course credit.

### *3.1.2. Stimuli*

All words were from four to six letters in length. Four hundred and eighty pairs were selected so that, as measured by the University of South Florida association norms (Nelson et al., 1998), all associations were unidirectional (no backward association occurrences) with a minimum forward strength of 0.05 (mean = 0.16). All words were presented in upper case, and

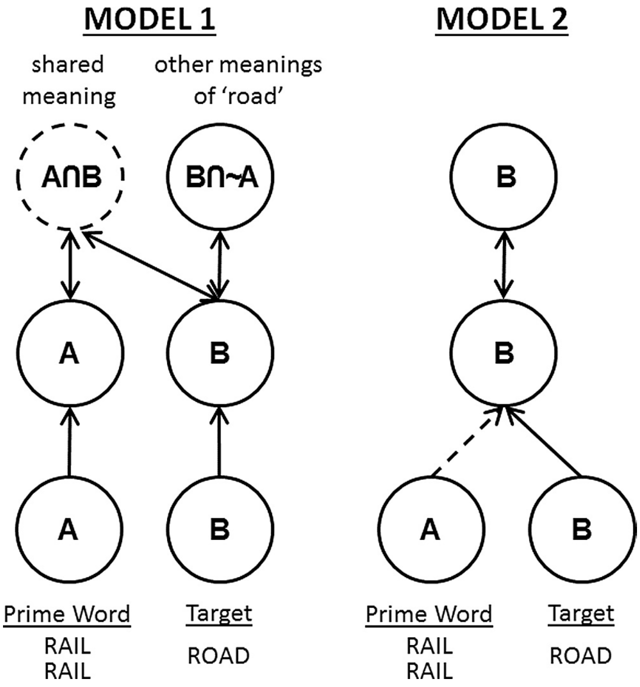


**Fig. 4.** The two models fit to the results of Experiment 2 with associative/semantic priming. MODEL 1 was used by Huber (2008) to explain priming advantages in the both associatively primed condition, in contrast to both primed deficits with repetition priming. In keeping with the assumed nROUSE hierarchy (see Fig. 2 labels), this model assumes that semantically related words have a shared representation at the top layer but not the middle layer. The semantic representation shared between a prime (A) and target (B) is shown as their intersection in mathematical set notation, as seen in the dashed circle. The representational size of the shared representation is the key free parameter in MODEL 1. Thus, if this parameter is set to 30% semantic overlap between prime and target, then the semantic representation that is unique to the target (indicated by the node labeled as the intersection of B with not-A) is set to the complement of this value (70%). MODEL 2 violates the assumed hierarchy by assuming that a prime and a semantically related target contact the same representation at layer 2. Thus, the key free parameter for this model is the connection strength from the prime to the target's layer 2 representation, akin to the model used for repetition priming of inverted words in Experiment 1. Because the other model parameters were set to values that best-fit Experiment 1, MODEL 2 necessarily produces a transition from positive to negative priming.

flanked by '#', so that the entire string was eight characters long. Table 2 presents example stimuli of each type for each prime condition.

### *3.1.3. Procedure*

Except as noted, all procedures were identical to Experiment 1. Because the linguistic properties of forward primes are not necessarily the same as the linguistic properties of backward primes, all trials used a target and foil pair of words that were either both chosen from the pool of backward primes or both chosen from the pool of forward primes. Furthermore, these words were selected to have the same number of letters. Conditions were presented in random order and no word appeared on more than one trial over the course of the entire experiment.

As in Experiment 1, the experiment began with practice trials followed by a block of trials to determine an appropriate target duration. For these trials, primes were unrelated to the choice words (neither primed). Words for these trials were selected by using three associated pairs to provide the words for two trials. For instance, consider the words pairs 1F/1B, 2F/2B, and 3F/3B, where the number refers to the associated word pair and F versus B indicates a forward versus backward relationship (i.e., the Backward word is an associate of the Forward word, but not vice versa). Thus, one neither primed trial was created by presenting 1F as the prime followed by 2B and 3B as the choices. A second neither primed trial was created by presenting 1B as the prime followed by 2F and 3F as the choices. The trials occurred in random order making it unlikely that these two trials were adjacent within the sequence of trials. Unlike Experiment 1, which used a prime duration of 500 ms during the threshold block of trials, Experiment 2 used a prime duration of 150 ms during the threshold block of trials. Otherwise, all procedures during the practice and threshold block of trials were the same as in Experiment 1.

After the threshold block of trials, the main experimental trials began and every trial primed either the target or the foil. In the forward association condition, the association was from the prime to the primed choice (e.g., 'saber' priming 'tooth'). The association was from choice word to the prime for the backward association condition (e.g., 'tooth' priming 'saber'). Trials were created using three associated word pairs to provide one forward priming trial and one backward priming trial. For example, the forward priming trial presented 1F as the prime followed by 1B and 2B as the choices. The backward priming trial presented 3B as the prime followed by 2F and 3F as the choices. There were 12 trials of each of the 20 conditions

**Table 2**
Conditions and example stimuli for Experiment 2.

| Direction | Primed choice | Example prime | Example target | Example foil |
|---|---|---|---|---|
| Forward | Target | RAIL | ROAD | BONE |
| Forward | Foil | PORCH | TEETH | LIGHT |
| Backward | Target | WORD | CODE | SEND |
| Backward | Foil | WELL | FINAL | ALIVE |

(5 prime durations, 2 primed choice, and 2 prime directions), for a total of 240 trials in the main experiment, broken into three blocks of 80.

## 3.2. Results

### 3.2.1. Priming effects

Average accuracy for each condition in Experiment 2 is shown in the left-hand graphs in Fig. 5 and the right-hand graphs show the corresponding priming effects (target primed minus foil primed). As seem in the figure, it appears that forward priming was slightly larger than backward priming, particularly following a 150 ms prime, but the priming effect was not reliably greater for forward priming than backward priming ($F(1,72) = 3.62$, $p = 0.061$). Furthermore, the three way interaction between primed choice, prime duration, and associative direction ($F(4,228) = 1.52$, $p = 0.197$) was not significant, supporting the conclusion that priming direction did not matter. Collapsing across forward versus backward priming, there was an interaction between prime duration and primed choice, ($F(4,228) = 9.05$, $p < 0.001$, $\eta^2 = 0.014$), with the magnitude of the priming effect increasing and then decreasing with increasing prime duration: The priming effect increased from the 17 ms to the 50 ms prime durations ($t(145) = -3.35$, $p_{holm} = 0.003$), did not change from the 50 ms to the 150 ms prime durations ($t(145) = -2.05$, $p_{holm} = 0.084$), decreased from the 150 ms to the 400 ms prime durations ($t(145) = 3.56$, $p_{holm} = 0.002$), and did not change from the 400 ms to the 2000 ms prime durations ($t(145) = 0.95$, $p_{holm} = 0.344$). In summary, there were no reliable differences between forward versus backward priming and both types of priming increased and then decreased in magnitude with increasing prime duration, but failed to transition to negative priming.

### 3.2.2. Model comparison

MODEL 1 assumes that associatively related word pairs have a shared lexical-semantic representation but no shared orthography. Thus, because the prime and target do not share any connections (see Fig. 4), priming should fail to become negative even following long duration primes. Furthermore, this provides a conceptual explanation of why forward and backward priming produced similar results, considering that the orthography of both words connect to the shared semantic representation. MODEL 1 is contrasted with MODEL 2 in Fig. 4, which assumes that the layer 2 representation for the target is directly activated by the prime (as shown by the dashed arrow from the visual input of the prime). In other words, this model assumes that associative priming is qualitatively the same as repetition priming (compare MODEL 2 in Fig. 4 to MODEL 1 in Fig. 2), but differs quantitatively (i.e., rather than a connection strength of 1.0 as appropriate to upright words, the connection strength from prime to associatively related target is less than 1.0 as indicated by the arrow being dashed rather than solid). Because the connection between layers 2 and 3 is shared between prime and target for MODEL 2, a sufficiently long prime duration results in negative priming as it becomes difficult to re-activate the layer 3 representation of the target.

In applying MODELs 1 and 2 to the Experiment 2 results, the same parameters were used as for Experiment 1 to force the models to provide a consistent set of predictions across repetition and associative/semantic priming. There were only two free parameters: a parameter capturing the magnitude of associative priming (the proportion of shared meaning for MODEL 1 or the prime-target layer 1 to layer 2 connection strength for MODEL 2) and the decision noise constant (capturing any overall accuracy difference between the experiments when collapsing across conditions, but, as reported in Appendix A, the best-fitting value for Experiment 2 was nearly identical to the value for Experiment 1). These models were applied to all 20 of the conditions of Experiment 2 (i.e., although the models assumed no differences between forward and backward priming, they were nevertheless fit to the full data set including the breakdown of priming direction). The $G^2$ goodness of fit for MODEL 1 was 71.8 whereas it was 83.2 for MODEL 2, supporting the assumption that layer 2 represents orthography whereas layer 3 contains semantic representations. Beyond the quantitative misfit, MODEL 2 is qualitatively inaccurate as seen in Fig. 5, producing negative priming for longer prime durations.

The behavior of MODEL 2 requires additional comment considering that the observed behavior of this model was not merely a scaled down version of repetition priming. More specifically, as applied to repetition priming in Experiment 1, the priming effect became negative and remained negative with increasing prime duration, but as applied to associative priming in Experiment 2, the priming effect became negative with increasing prime duration but then tipped back toward positive priming for prime durations even longer than those used in the experiment. This occurred because the best-fitting parameters from Experiment 1 set the visual layer to have a slower time constant than the orthographic layer. As a result, layer 2 experienced full habituation first, before layer 1, but once layer 1 became fully habituated, layer 2 was able to recover somewhat because it was not driven as strongly by layer 1. This dynamic occurred for repetition priming in Experiment 1 (there is a hint of this in the model, with a slight lessening of negative priming), but it was not as pronounced considering

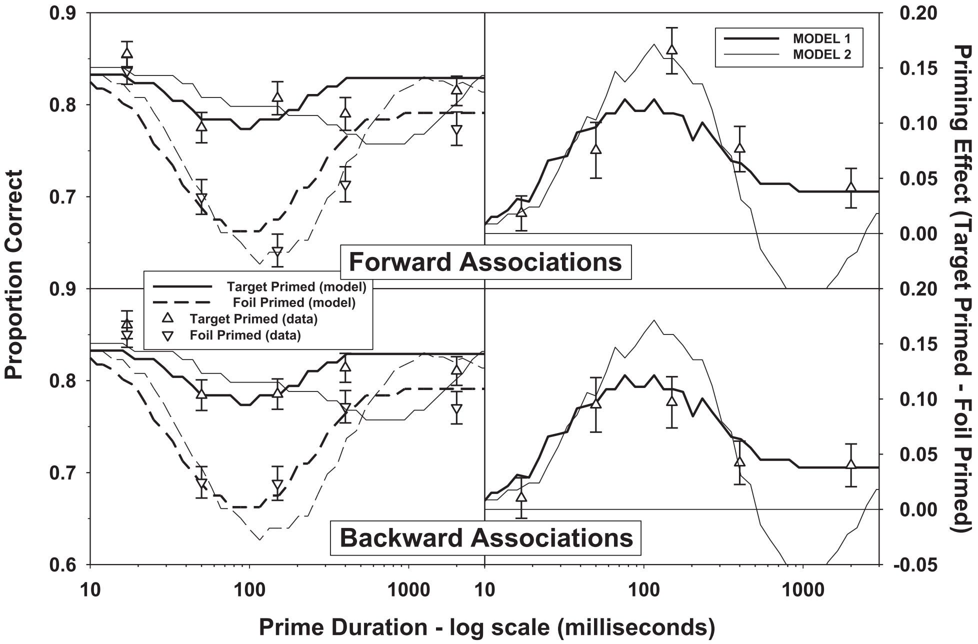


**Fig. 5.** Experiment 2 results (symbols) and model results (lines). The left set of graphs shows average accuracy in each condition and the right set of graphs shows priming effects by taking the difference between the target primed and foil primed conditions. Error bars are +/−1 standard error of the mean, or standard error of the mean difference. The 20 conditions of this experiment were fit using 2 free parameters for each model, with MODEL 1 producing a $G^2$ error of 71.8 versus 83.2 for MODEL 2. For both models, most of the parameters were set to the best-fitting values from Experiment 1, which was identical except for the use of repetition priming rather than associative priming. The two free parameters were decision noise, to capture performance differences for the subjects of Experiment 2 as compared to Experiment 1, and a parameter capturing the magnitude of associative priming (shared semantics for MODEL 1 and the layer 1–2 prime-target connection strength for MODEL 2). Both models assumed no difference between forward versus backward associations. As seen in the figure, MODEL 2 incorrectly predicted that priming should become negative following a 2000 ms prime. The seemingly noisy behavior of the model reflects the use of discrete time steps (1 ms), with the bumps and wiggles indicating discrete changes in the integer number of milliseconds for the target and/or foil perceptual identification latencies (this was not obviously in Fig. 3 because of the y-axis scale).

that repetition priming sets the layer 1 to 2 connection at 1.0 whereas for associative priming in MODEL 2 the best-fitting connection from the prime's layer 1 to the target's layer 2 was 0.3. In any event, there is no hint of negative priming in the observed results, in contradiction to MODEL 2.

### 3.3. Discussion

As a function of increasing prime duration, associative priming increased and then decreased, but remained positive. This failure to produce negative priming was predicted by the nROUSE model considering that the orthographic to semantic connections of the target word are not activated by a semantically related prime (i.e., there is no habituation for these connections because although the prime and target share meaning, they do not share orthography). Instead, habituation only reduces the amount of lingering semantic activation (i.e., it only reduces the semantic priming benefit). Huber et al. (2001) examined two different prime durations in a between-subjects design, finding a reduction in the magnitude of associative priming but no reversal to negative priming. However, prior to the current experiment, the full nonlinear time course of associative priming was never examined with prime duration as a within-subjects factor. This is a key confirmation of the nROUSE model's predictions. Experiment 2 also confirmed the prediction that associatively asymmetric word pairs produce similar priming dynamics and magnitude regardless of the direction of priming (e.g., regardless of whether 'saber' is shown as the prime and 'tooth' as the target, or vice versa). This prediction follows from the assumption that associative priming reflects semantic content shared between prime and target.

Using parameters that explained the transition to negative priming with repetition priming in Experiment 1, MODEL 1 accurately captured the observed associative priming results. In contrast, MODEL 2 attempted to explain associative priming in the same manner as repetition priming, producing negative priming that was not seen in the data. In summary, the observed differences between Experiments 1 and 2, and the model comparison results, support the assumption that layer 2 of the nROUSE model represents orthography and layer 3 contains semantic representations.

One concern when comparing Experiment 2 and Experiment 1 is that the subjects in each experiment may have adopted different strategies in regard to the primes considering that all of the priming conditions were repetition priming in Experiment 1 whereas all of the priming conditions were associative priming in Experiment 2. To address this limitation, Experiment 3 examined both of types of priming in a within-subjects design.

## 4. Experiment 3

Aside from using a within-subjects design to examine both repetition and semantic priming, Experiment 3 generalized the results of Experiment 2 by defining semantic priming through natural language co-occurrence statistics rather than word association norms. Many successful models of semantics use corpus co-occurrence statistics as the base measure to create semantic representations (Griffiths, Steyvers, & Tenenbaum, 2007; Jones & Mewhort, 2007; Landauer & Dumais, 1997; Lund & Burgess, 1996) and defining semantic relations in this manner allowed us to use a corpus to create conditions that combined the effects of repetition and semantic priming.

Rather than two copies of an upper case prime word, Experiment 3 presented a single lower-case prime phrase that included a prime word followed by a closed class word (e.g., 'stake-a') to evoke a specific target word (e.g., 'CLAIM') that might complete an idiom phrase as defined by natural language co-occurrence statistics. We termed this condition the 'expected non-repetition' condition and we anticipated that this condition would be similar to associative priming in Experiment 2. Experiment 3 also included a 'non-expected repetition' condition (e.g., 'cloth-of' priming 'CLOTH'), which should be similar to the word repetition priming conditions of Experiment 1. Finally, Experiment 3 combined repetition and semantic priming by examining whether likely repetitions (e.g., 'back-to' priming 'BACK') produce a different pattern than unexpected repetitions. We termed this condition the 'expected repetition' condition. The key difference between the expected and non-expected repetition conditions is that for the expected repetition condition, the appearance of the closed class word in combination with the prime word may elicit a specific meaning of the target word (e.g., the meaning of 'back-to-back') that is not otherwise evoked by the prime word in isolation. Thus, we anticipated that the expected repetition condition would reflect the combined effects of the non-expected repetition conditions and the expected non-repetition conditions.

It is important to note that our use of the term 'expected' refers to natural language statistics (i.e., things learned prior to the start of the experiment) rather than whether primes are more or less useful for determining the correct response. An expectation that priming will occur can be varied through the proportion of trials that present a related target (e.g., Bodner & Masson, 2003; Weidemann et al., 2008). This makes primes more or less diagnostic and can produce effects even when people are unaware of these manipulations. However, as in Experiments 1 and 2, the primes in Experiment 3 were just as likely to indicate the incorrect choice as the correct choice (i.e., the primes were non-diagnostic for all three types of priming). The question of interest was whether expectations from everyday language would produce results similar to associative priming and whether these effects would be additive with repetition priming.

Model 1 in Fig. 6 outlines nROUSE's preferred architecture when combining semantic and repetition priming. First, consider the non-expected repetition condition. In this case, the isolated prime word and target have the same meaning by virtue of being the same word, but the three-word phrase (e.g., cloth of cloth) does not evoke any meaning beyond the typical meaning of the repeated word. In Fig. 6, this corresponds to removing the dashed circle because there is no shared meaning between the closed class word and the target. Without this shared meaning, the model reduces to the basic repetition priming model (e.g., MODEL 1 from Experiment 1). Next, consider the expected non-repetition condition (e.g., stake-a priming CLAIM). In this case, the prime word is orthographically different than the target and the connection inside the gray box is removed. Without this connection, the model reduces to the basic associative/semantic priming model (e.g., MODEL 1 from Experiment 2). Finally, consider the expected repetition condition (e.g., back-to priming BACK). In this case, the target word is both repetition primed by the prime word (the B' to B connection) and semantically primed by the closed class word (the intersection of A and B in the dashed circle). Because performance in nROUSE is dictated by the summation of all of the semantic representations for each choice word, the priming effects in the expected repetition condition should be similar to the priming effects of the non-expected repetition condition added with the priming effects from the expected non-repetition condition.

Finally, this experiment provided a great deal of constraint on the nROUSE model by also including a baseline neutral prime condition with a prime that was unrelated to either choice (i.e., the 'neither primed' condition). The model makes specific predictions regarding the relative placement of the neither primed condition as compared to the target primed and foil primed conditions. More specifically, if repetition priming entails a depleted orthographic to semantic connection, there should be a true perceptual deficit for repetition priming (but not necessary semantic priming) such that the average of the target primed and foil primed conditions is lower than the neither primed condition.

### *4.1. Method*

#### *4.1.1. Participants*

A total of 119 undergraduate psychology students participated in this experiment for course credit.

#### *4.1.2. Stimuli*

Three-word idiom phrases were selected based on an initial list of 1560 four to six letter words (hereafter, the content word list), obtained from the MRC linguistic database (Coltheart, 1981), all with Kucera and Francis (1967) written word frequency above 10. Additional one to four letter closed-class words were selected to be used as the second word in the two-word prime phrases. The closed class words were: a, an, and, are, as, at, by, for, from, go, he, in, is, it, my, no, of, on,

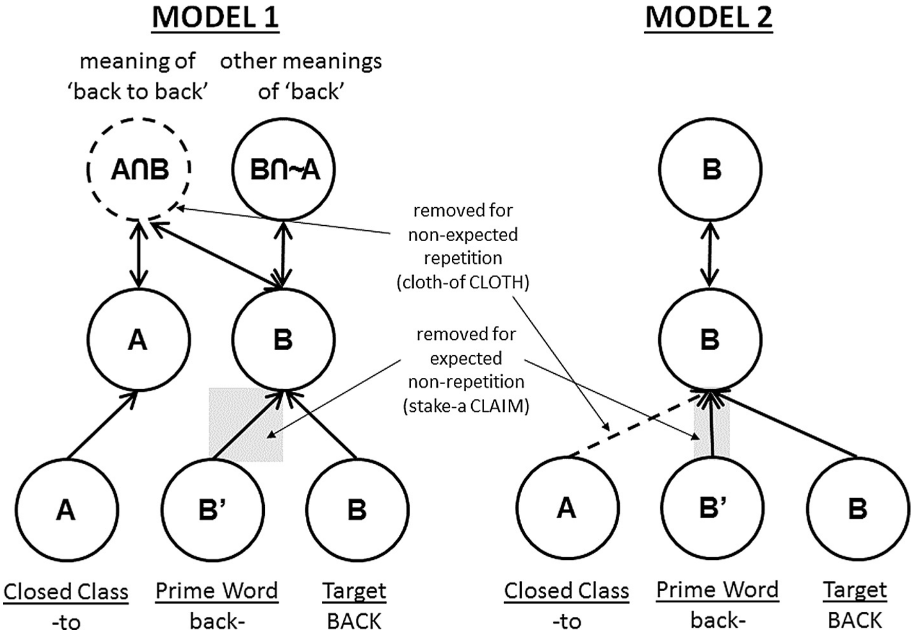


**Fig. 6.** The two models fit to the results of Experiment 3. Because the prime phrase contains two different words (e.g., 'back' and 'to'), two separate visual inputs are assumed. For the non-expected repetition condition, MODEL 1 reduces to the MODEL 1 used in Experiment 1 for repetition priming because there is no meaning associated with the three-word idiom aside from the meaning of the prime word in isolation (e.g., 'cloth-of-cloth' does not mean anything). For the expected non-repetition condition, MODEL 1 reduces to the MODEL 1 used in Experiment 2 for associative priming because the visual response to the prime word is not connected to the target word's orthography. Except for the small role of semantic feedback, these two forms of priming are additive at the layer 3 representation and the priming effect for the expected repetition condition was predicted to be the summation of the priming effects for the other two conditions. This can be contrasted with MODEL 2, which adopts the same assumption as the MODEL 2 in Experiment 2, such that repetition priming from the prime word's visual response and semantic priming from the closed class word's visual response are combined in the layer 2 representation. Because choice behavior is dictated by layer 3, but the two forms of priming are combined in layer 2, repetition and semantic priming are sub-additive in MODEL 2.

one, or, so, that, the, them, this, to, up, was, we, and with. The Google-n-gram database (Brants & Franz, 2006), which tallies the number of online occurrences of word sequences, was used to determine the frequency of each possible three-word idiom phrase created by a first content word, a closed class word, and final content word, allowing for repeats of the content word. For each possible three-word idiom phrase, we calculated the conditional probability of the final content word given the two-word prime phrase (the proportion of times the third word occurred after the first two words out of all occurrences of the first two words). This conditional probability is commonly referred to as the transition probability. However, a high transitional probability can occur even with very rare (i.e., not well-learned) idiom phrases if the first two words of the triple rarely co-occur. Therefore, we also considered triplet frequency in calculating an 'expectancy score' to select for idiom phrases that were both fairly common and highly predictive of the final word given the two-word prime phrase. The triplet frequencies (log transformed) and transitional probabilities were individually z-scored relative to all word triplets. The expectancy score was twice the z-scored transitional probability plus the z-scored log-triplet frequency (i.e., more emphasis on transitional probability than on frequency).

Potential three-word idiom phrases for the expected repetition and expected non-repetition conditions had expectancy scores greater than three (the top 0.25% of triplets). Content words only appeared in one of the triplets and so an ordered selection process was used to maximize the number of available stimuli for each condition. The order of selection for triplets that met these requirements was expected repetition, then expected non-repetition, then non-expected repetition. The transition probability for the non-expected repetition condition was zero (e.g. 'cloth' never followed 'cloth-of' in the n-gram database). The final triplets for each condition were selected by hand. For each word triplet, a matched choice word was chosen from the remaining content words with the constraints that the choice word contain the same number of letters as the third word of the triplet and that the choice word have a transitional probability of zero based on the first two words of the triplet (e.g. for the triplet 'crowd-go wild', the word 'clue' was assigned as an alternative choice word because 'crowd-go clue' never occurred). This resulted in 90 word triplets for each of the three types of priming, each with an alternative third word. An additional 48 expected non-repetition triplets were selected for use in the threshold trials. An additional 20 triplets were used for the practice trials, representing a mix of conditions.

### 4.1.3. Procedure

Except as noted all procedures were identical to Experiment 2. The prime phrase was centered and in lower case, with the content word and closed class word separated by a hyphen (see Table 3 for examples). Targets and choices were presented in upper case to distinguish them from primes. There were 20 practice trials, followed by 72 threshold trials (16 at each of the 4 target durations). During the threshold trials, there was an equal mix of neither primed, target primed and foil primed conditions, with all priming of the expected non-repetition type. The main experiment consisted of three blocks of 90 trials. The 90 trials in each block consisted of (in random order) all combinations of the three types of priming (expected repetition, expected non-repetition, and non-expected repetition), the three primed choice possibilities (neither, target, or foil primed), the five prime durations, and the two possible screen positions of the target (left versus right).

For target primed conditions, the expected and/or repeated third word in the triplet was presented as the target and the alternative choice word the foil. So ‘crowd-go wild’ as a target primed condition would present ‘crowd-go’ as a prime in the center of the screen, ‘WILD’ as a target also in the center and both ‘WILD’ and ‘CLUE’ as choices on the left and right sides of the screen. For foil primed conditions, the expected and/or repeated third word was the foil choice and the alternative choice word was the target. Neither primed trials were created by swapping choice words (the third word and its alternative) between a specific pair of triplets within conditions. For example, the ‘crowd-go wild/clue’ triplet was paired with ‘kiss-of death/south’ triplet, resulting in stimuli for two neither primed condition trials, one with a prime of ‘crowd-go’, a target of ‘DEATH’ and choices of ‘DEATH’ and ‘SOUTH’, and the other with a prime of ‘kiss-of’, a target of ‘WILD’ and the choices ‘WILD’ and ‘CLUE’. Triplets were paired with the constraint that the transitional probabilities were zero.

For each participant, each pair of word triplets was randomly assigned to one of the three primed choice conditions (neither, target, or foil primed), subject to the constraint that an equal number of trials occur for each condition. This procedure ensured that the same targets appeared with the same foils while randomly varying triplets between primed choice conditions. Thus, there were separate neither primed conditions for each of the three types of priming to control for potential word differences between the three types of priming. This was an important control considering that the target/foil words in the expected repetition condition were of higher frequency than the target/foil words in the non-expected repetition condition: mean log frequency of expected repetitions 17.15 versus 16.25 for non-expected repetitions, $t(166) = 5.09$, $p_{holm} < 0.001$. As seen in Fig. 4, the three neither primed conditions were nearly identical, suggesting that these frequency differences did not affect performance. Furthermore, based on the results of Experiment 1, a substantial frequency difference should have increased the rate of transition from positive to negative priming for the expected repetition condition as compared to the non-expected repetition condition, although as seen in Fig. 4, the opposite was true. As further verification, we performed a median split of the expected repetition condition based on frequency of the prime words. Nevertheless, the priming results were nearly identical across these two frequency levels.

## 4.2. Results

### 4.2.1. Priming effects

Average accuracy for all conditions of Experiment 3, including the neither primed conditions, is shown in left-hand graphs of Fig. 7, and the right-hand graphs show the corresponding priming effects (target primed minus foil primed). There was a significant three way interaction between priming type (non-expected repetition, expected non-repetition, or expected repetition), primed choice (target, foil, or neither), and prime duration ($F(16,1888) = 9.82$, $p < 0.001$, $\eta^2 = 0.020$). Replicating the previous repetition priming results, when targets were non-expected repetitions of prime words (e.g. ‘cloth-of’ priming ‘CLOTH’), the priming effect initially grew and then shrank with increasing prime duration. Target primed accuracy was greater than foil primed accuracy for prime durations from 17 and 150 ms ($t(118) \geq 6.70$, $p_{holm} < 0.001$). When correcting for multiple comparisons, it cannot be concluded that negative priming occurred for the non-expected repetition condition for a 2000 ms prime ($t(118) = -2.27$, $p_{holm} \leq 0.125$). However, we note that this was an a priori prediction rather than a post hoc comparison, and the uncorrected type one error rate is 0.0125. For the expected repetition condition (e.g., ‘walk-the’ priming ‘WALK’), there was positive priming for prime durations from 17 and 150 ms ($t(118) \geq 10.08$, $p_{holm} < 0.001$), but, unlike the non-expected repetition condition, there was no hint of negative priming following a 2000 ms prime ($t(118) = -0.98$, $p_{holm} = 0.977$). A separate repeated measures ANOVA was run to directly compare the expected and non-expected repetition conditions. There was a significant interaction between type of priming (expected versus non-expected) and primed choice ($F(2,236) = 6.25$, $p_{holm} = 0.001$, $\eta^2 = 0.020$), revealed as an overall greater priming effect for expected repetitions when collapsing across prime duration ($t(1188) = -2.33$, $p_{holm} = 0.020$). Finally, for the expected non-repetition priming condition (e.g. ‘crowd-go’ priming ‘WILD’), there was a positive priming effect for prime durations from 50 to 400 ms ($t(118) = 4.95$, $p_{holm} < 0.001$). For a 2000 ms prime, when correcting for multiple comparisons, there was marginally significant positive priming effect ($t(118) = 2.63$, $p_{holm} = 0.058$).

### 4.2.2. Model comparison

MODEL 1 was contrasted with the same alternative MODEL 2 as used in Experiment 2. For MODEL 2 (see Fig. 6), semantic priming was assumed to arise from a direct connection between the closed class word and the target’s layer 2 representation. Thus, rather than combining semantic and repetition priming at layer 3 as in MODEL 1, these two forms of priming were combined in layer 2. Because this addition occurred prior to the layer 2–3 connection, semantic priming was predicted to potentially become negative with longer prime durations. Furthermore, the expected repetition condition was predicted

**Table 3**
Conditions and example stimuli for Experiment 3.

| Type of priming | Primed choice | Example prime | Example target | Example foil |
|---|---|---|---|---|
| Non-expected repetition | Target | cloth-of | `CLOTH` | `FERRY` |
| Non-expected repetition | Foil | cloth-of | `FERRY` | `CLOTH` |
| Non-expected repetition | Neither | cloth-of | `WAIST` | `CHILL` |
| Expected non-repetition | Target | stake-a | `CLAIM` | `ALIEN` |
| Expected non-repetition | Foil | stake-a | `ALIEN` | `CLAIM` |
| Expected non-repetition | Neither | stake-a | `BRIDGE` | `SLEEVE` |
| Expected repetition | Target | back-to | `BACK` | `FIST` |
| Expected repetition | Foil | back-to | `FIST` | `BACK` |
| Expected repetition | Neither | back-to | `FIRE` | `BOLT` |

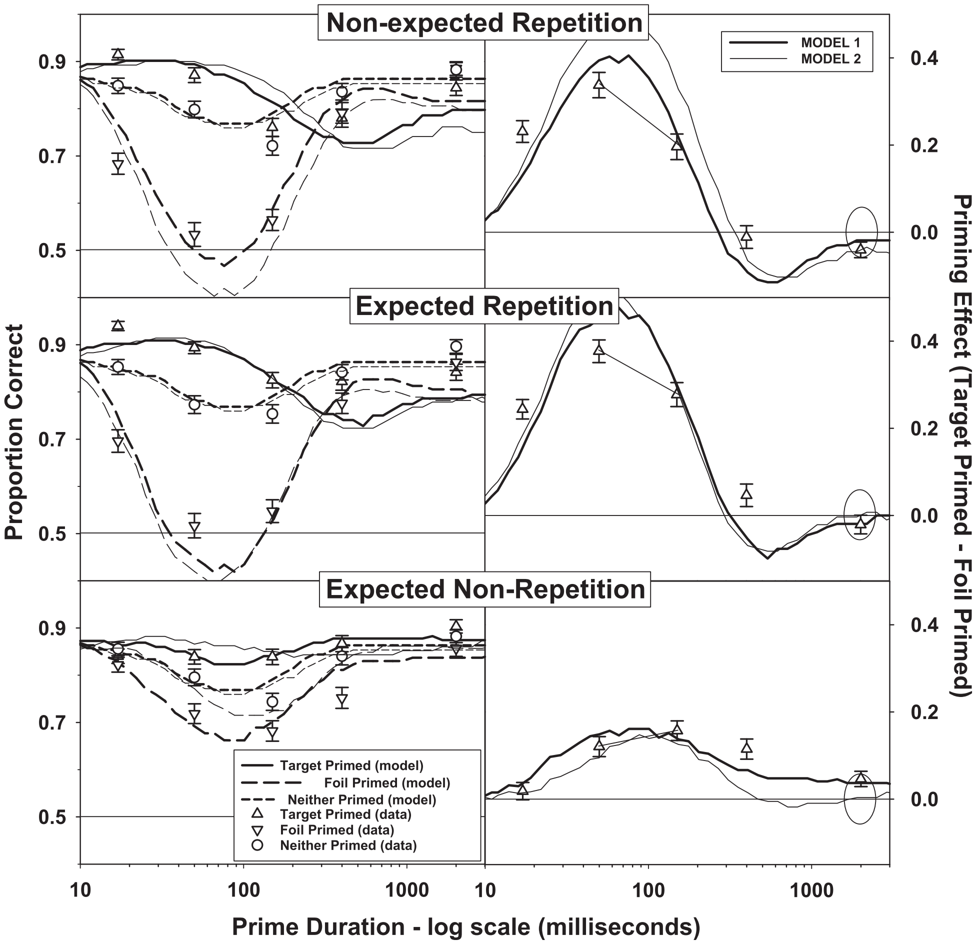


**Fig. 7.** Experiment 3 results (symbols) and model results (lines). The left set of graphs shows average accuracy in each condition and the right set of graphs shows priming effects by taking the difference between the target primed and foil primed conditions. Error bars are +/−1 standard error of the mean, or standard error of the mean difference. The 45 conditions of this experiment were fit using 3 free parameters for each model, with MODEL 1 producing a $G^2$ error of 273.7 versus 417.2 for MODEL 2. For the right-hand set of graphs, the lines connecting the 50 ms and 150 ms priming effects of the observed data highlight the finding that priming continued to increase for non-repetition (semantic) priming whereas priming decreased for the two kinds of repetition priming. The ovals at 2000 ms, which are centered at a priming effect of 0, highlight the changes in the direction of priming when comparing across the three kinds of priming. The observed changes in the priming dynamics and magnitudes were well explained by MODEL 1, which produced additive effects for the expected repetition condition as compared to the other two conditions. In contrast, MODEL 2 produced similar results for the two kinds of repetition priming.

to reflect a sub-additive combination of the non-expected repetition and the expected non-repetition conditions, owing to the non-linearity of the shared neural representations (i.e., there is an upper-bound for the voltage potential of the layer 2 representation).

The modeling of Experiment 3 was highly constrained, using just three free parameters to capture the 45 different conditions that included both semantic and repetition priming as well as the baseline neither primed condition. These free parameters were the decision noise parameter, to capture different accuracy levels for the subjects of Experiment 3 as compared to Experiment 1, a parameter capturing the magnitude of semantic priming (the proportion of shared semantic content for MODEL 1 and the strength of the layer 1 to 2 connection for MODEL 2), and a third parameter modulating the strength of the layer 1–2 connection from the prime in all conditions. This third parameter reflected the difference between using lower-case letters along the midline for the primes in this experiment rather than upper-case letters above and below the midline for Experiments 1 and 2. Sensibly, the best-fitting value for this third parameter was higher than the default of 1.0, suggesting that midline letters were more readily identified than ones presented above and below the midline.

Given the constrained nature of this application of nROUSE, it is unsurprising that the $G^2$ error was higher for MODEL 1 (273.7) as compared to the MODEL 1 results from Experiments 1 and 2. Crucially, the fit of MODEL 2 to these data was substantially worse (417.2) than MODEL 1. As seen in Fig. 7, MODEL 1 adequately captured all of the qualitative trends and provided a reasonable quantitative description of the data. Key qualitative predictions are highlighted by connecting the 50 ms and 150 ms priming effects and by centering ovals at the 2000 ms priming effect, similar to annotations provided for Experiment 1. The lines make it clear that for both kinds of repetition priming, the priming effect was already decreasing by 150 ms whereas for semantic priming, the priming effect was still increasing. Both models qualitatively captured these trends, although the dynamics of MODEL 1 more closely tracked the observed data. The priming effects at 2000 ms, as highlighted by the ovals, highlight a key differences between the models. The observed data show negative priming for the non-expected repetition condition, positive priming for the expected non-repetition condition, and a lack of priming for the expected non-repetition condition, with the latter reflecting the additive effects of repetition and semantic priming according to MODEL 1, which captured all of these effects. In contrast, MODEL 2 produced negative priming for the non-expected repetition condition and a lack of priming for the other two conditions. MODEL 2 misfit the data in other important aspects: unlike MODEL 1, MODEL 2 produced very similar priming effects for both kinds of repetitions whereas MODEL 1 correctly produced larger priming for expected repetitions as compared to non-expected repetitions. Also, as seen in the left-hand set of graphs, MODEL 1 correctly captured the large deficit for the average of the target primed and foil primed conditions, as compared to the neither primed condition in the case of repetition priming, whereas this was not true for semantic priming (e.g., compare the two models ability to capture the target, foil, and neither primed conditions following a 50 ms semantically related prime, as seen in the lower left-hand graph).

### 4.3. Discussion

Experiments 1 and 2 produced different priming dynamics for repetition versus associative priming, but it is possible that these differences reflected some kind of different response strategy considering that this was a between-subjects comparison. Experiment 3 replicated both of these effects using a within-subjects design, finding a different time course for semantic priming as compared to repetition priming, with repetition priming eventually becoming negative at the longest prime duration whereas semantic priming was still positive at this same duration. Furthermore, Experiment 3 used a different metric for defining semantic relations based on natural language co-occurrence statistics, which are commonly used to build computational models of semantic representations. Finally, comparing words that were expected to repeat versus words that were not expected to repeat, Experiment 3 confirmed the additive nature of repetition and semantic priming. The observed data and model comparison results provide further support for the representational assumptions of the nROUSE model.

Beyond combining the repetition priming results of Experiment 1 and the associative priming results of Experiment 2, Experiment 3 included a baseline neither primed condition. This confirmed that these priming effects are a perceptual bias, with the neither primed condition somewhere between the target and foil primed conditions regardless of prime duration or type of priming (i.e., priming produced a pattern of costs and benefits depending on whether the correct or incorrect choice was primed). Similar to the results of Huber (2008), the neither primed condition was above the average of the target and foil primed conditions with repetition priming, suggesting a perceptual deficit with repetition priming, whereas the neither primed condition was slightly below the average of the target and foil primed conditions with expected non-repetition priming, suggesting a perceptual gain with semantic priming. According to nROUSE, a perceptual deficit occurs with repetition priming because perception of the briefly flashed target is sluggish owing to habituated orthography. In contrast, there is no habituation for the orthography of the target in the expected non-repetition condition and so the briefly flashed target in this condition is perceived more efficiently owing to top-down support from the shared meaning between prime and target.

The neither primed condition also revealed a u-shaped pattern as a function of prime duration, with the lowest accuracy level following a 150 ms prime. In other words, an unrelated prime word produces u-shaped forward masking as a function of prime duration. According to the nROUSE model, this u-shaped forward masking is directly related to the semantic priming results. More specifically, 150 ms is the prime duration that produced the greatest lexical-semantic activation for the meaning of the prime phrase and this provided the greatest inhibition for the meaning of an unrelated word in the neither primed condition as well as the greatest top-down support for the orthography of a related target in the expected non-repetition condition.

## 5. General discussion

How does perception accurately identify the currently viewed object without confusion from recently viewed objects? In an RSVP priming task, words are presented in rapid succession, faster than reading speed, which makes accurate identification of a target word difficult owing to lingering perception of the letters and/or meaning of an immediately preceding prime word. This produces a perceptual bias, resulting in better performance if the correct word is primed but worse performance if the incorrect word is primed. For prime durations that are similar to reading speed (e.g., between 150 and 400 ms), the bias to confuse the prime with the target is eliminated. For even longer prime durations (e.g., between 400 and 2000 ms), a reversed bias occurs for repetition priming. In this case, a perceptual deficit boosts performance if the incorrect word is primed but harms performance if the target is primed. This biphasic pattern of results as a function of prime duration was first demonstrated by Huber et al. (2001), and these findings were explained in terms of evidence discounting in an approximately optimal Bayesian perceptual decision process (the ROUSE model). A subsequent study by Huber and O'Reilly (2003) proposed that neural habituation is the mechanism that produces evidence discounting, and the dynamic time course of priming was explained with a dynamic neural network model (nROUSE) that contained habituation dynamics.

According to the nROUSE model, habituation exists to reduce perceptual source confusion and is applied at all levels of processing (e.g., from perception of simple line segments to perception of meaning). For repetition priming of words, habituation of the connections between orthographic and lexical/semantic representations is crucial because these are the connections that are reused when a word repetition appears in a different screen location as a target word or a choice word. Habituation of these connections clears the way to perceive a novel subsequent word with minimal orthographic source confusion and it reduces the amount of semantic carryover from one word to the next. However, this mechanism for parsing the sequence of words comes at a cost, producing deficits when the orthography of the previous word repeats.

The core assumption of the nROUSE model is the dynamics of habituation, and more specifically synaptic depression. These dynamics predict that stimuli with weaker connections (e.g., a non-word) should produce a perceptual response that while initially weak (i.e., worse performance in the threshold identification task), is a response that exhibits a slower habituation time course (i.e., a slower crossover to negative priming). Experiment 1 confirmed this prediction by comparing different kinds of stimuli that were expected to differ in connection strength at different levels in the representational hierarchy (non-words were expected to have a weak connection from orthographic to lexical-semantic representations and inverted words were expected to have a weak connection form visual to orthographic representations). Critically, these predictions were made based on the original representation assumed by the nROUSE model (MODEL 1) and a plausible alternative representation (MODEL 2) produced a quantitatively and qualitatively worse fit to the data. Experiment 2 tested the representational hierarchy by examining associative priming, confirming the prediction that associative priming should rise and fall, but never become negative with increasing prime duration. In contrast, an alternative model assuming that associative priming is a weaker form of repetition priming incorrectly predicted that associative priming would become negative with increasing prime duration. Finally, Experiment 3 combined these effects in a within-subjects design, using a two-word prime phrase followed by a target word that might or might not be expected based on natural language co-occurrence statistics. The results of Experiments 1 and 2 were replicated in this experiment and repetition and semantic priming were found to be additive as expected with the assumed representational hierarchy.

### 5.1. Comparison to other theories of perception and priming

The center-surround theory of Dagenbach, Carr, and Wilhelmsen (1989) proposes that with limited perceptual information, the available information receives attentional enhancement (i.e., center enhancement), which produces repetition priming benefits in a lexical decision task with masked priming, while, at the same time, related information is suppressed (i.e., surround inhibition), which produces semantic priming deficits. Thus, center-surround theory makes opposite predictions to the nROUSE model when comparing repetition and semantic priming. Carr and Dagenbach (1990) found support for this theory, finding a priming benefit for repetition priming and a trend toward a priming deficit for semantic priming using the same prime duration. Although the semantic priming deficit was not statistically reliable in the Carr and Dagenbach study, Kahan (2000) successfully replicated this result. Kahan hypothesized that these lexical decision priming deficits reflect post-perceptual decision strategies related to the expected type of priming (e.g., the reader is slowed down by attempting to retrospectively determine the identity of the near-threshold prime based on the identity of the target). Supporting this hypothesis, Kahan found that naming latencies replicated the finding of semantic priming deficits combined with repetition priming benefits in a context of mostly semantic priming trials but this pattern reversed to semantic priming benefits combined with repetition priming deficits in a context of mostly repetition priming trials. The current study attempted to minimize the role of decision strategies by using an equal mix of foil primed and target primed trials and by providing trial-by-trial accuracy feedback to make it clear that choices based on the prime were just as likely to hurt as help. Supporting the claim that decision strategies did not contribute to the current experiments, the nROUSE model captured the results of all three experiments even though each experiment used a different mix of priming types: Experiment 1 used only repetition priming; Experiment 3 used a mix of repetition and semantic priming; and Experiment 2 used only semantic priming.

The processing dynamics in the nROUSE model are similar to the mismatch theory of Johnston and Hawley (1994) and the temporal integration theory of Grainger and Jacobs (1999). These theories have not been applied to the forced-choice perceptual identification task used in the current study, but if they were, they might make similar predictions. Nevertheless, we highlight some important differences. The integration theory of Grainger and Jacobs is not implemented as a computational model, but similar to the activity dependent synaptic depression implemented in nROUSE, this theory proposes a dynamic reset mechanism through self-inhibition. However, the reset mechanism of integration theory is only applied to orthographic representations whereas the nROUSE model is a general theory of the perceptual dynamics, with habituation occurring for all representations. For instance, nROUSE explained the difference between upright and inverted words by assuming a weakened connection between the visual input and orthography for inverted words (i.e., a change in the visual reset dynamics). However, an explanation at the level of visual reset dynamics is outside the scope of integration theory.

Because synaptic depression is found throughout the cortex (Thomson & West, 1993), nROUSE is generally applicable to other tasks that involve temporal separation, such as 'repetition blindness' (Hochhaus & Johnston, 1996; Kanwisher, 1987; Masson, 2004), the 'attentional blink' (Raymond, Shapiro, & Arnell, 1992; Rusconi & Huber, submitted for publication), and 'inhibition of return' (Dukewich, 2009; Posner & Cohen, 1984; Rieth & Huber, 2013), regardless of whether linguistic or non-linguistic stimuli are used. Similar to nROUSE, the mismatch theory of Johnston and Hawley (1994) holds the same promise of providing a unified account of these RSVP tasks. However, a computational model of this theory has only been applied to the visual search task and the specific finding that repetition priming of distractors causes visual pop-out for a novel target (see Davelaar et al. (2011) for a closely related finding and application of nROUSE to this finding). More generally, it is not clear how to extend mismatch theory to the current results or other identification paradigms considering that it requires specialized 'iconic nodes' that receive top-down inhibition from separate identity and location nodes. In contrast to mismatch theory, the dynamics of nROUSE can be applied to the representations of any existing theory, without requiring iconic nodes. This is achieved by including dynamic self-suppression based on the known properties of synapses.

In terms of Marr's (1982) different levels of explanation, nROUSE is an implementation-level model (i.e., a neural model) of the computational-level theory that recently viewed perceptual objects provide a competing cause of the current percept. According to this theory, the properties of recently viewed objects should be 'explained away' or 'discounted', and the ideal amount of discounting can be calculated using Bayes rule. This provides a causal explanation for the habituation dynamics of the nROUSE model. From this perspective, what is the computational-level explanation of the reported priming time course differences for different kinds of stimuli? According to Anderson's (1991) 'Rational Analysis' as applied to the current results, differences in the magnitude of prime discounting when comparing different types of words and word pairs can be understood in terms of the statistical regularities of natural language (see also Norris & Kinoshita, 2008). For instance, based on everyday experience, perception of a particular letter is more likely the result of a high frequency word as opposed to a once-in-a-lifetime non-word. Because a non-word is an unlikely cause of the percept, a rational/normative reading process should apply less discounting (and thus less negative priming) to the letters of a non-word prime.

## 6. Conclusions

The properties of recently viewed objects should be discounted to clear the way for accurate identification of subsequent objects with minimal source confusion. However, when objects are presented in rapid succession, such as in an RSVP task, there is inadequate time to perceive and discount previously viewed objects. A neural model of this theory, nROUSE, supposes that habituation through synaptic depression is the brain's trick for producing this discounting. Yet this trick comes with cost, producing repetition deficits. This model has successfully explained the time course of priming and temporal separation in word priming and a variety of other tasks but the key prediction that stronger connection strengths should produce faster habituation was never previously tested. As predicted by differences in connection strength, repetition priming of words produced a faster transition to negative priming as compared to repetition priming of non-words or inverted words. These predictions were based on the assumed three-layer hierarchy of the model, and as predicted with this hierarchy, increases in prime duration caused a transition from positive to negative priming for repetition priming, but only a decrease in the magnitude of positive priming for associative priming. Finally, these repetition priming and semantic/associative priming effects were combined, using three-word idioms in which the first two words (e.g., 'back-to') served as a prime phrase before presentation of the subsequent target word (e.g., 'BACK').

These results have important implications for the use of word priming in the study of psycholinguistics (e.g., the magnitude and even the direction of priming depends on psychophysical factors, such as prime duration, on psycholinguistic factors, such as natural language frequency, and on complex interactions between these factors.). Beyond this specific application of the model, confirmation of the predicted relationship between connection strength the rate of habituation, supports the general claim that neural habituation exists to solve a temporal parsing problem, with implications for any perceptual task in which objects are presented in rapid succession.

## Author note

This research was a portion of Cory Rieth's doctoral dissertation. Cory Rieth was supported by NSF BCS-0843773 to DE Huber, NSF IGERT Grant DGE-0333451 to GW Cottrell/VR de Sa, and the Rita L. Atkinson Graduate Fellowship. David Huber was supported by NSF BCS-0843773 and NSF BCS-1431147.

## Appendix A. nROUSE mathematical details

The reported simulations used a version of nROUSE that was slightly changed from the original model to reflect: (1) an updated understanding of synaptic depression and (2) a stochastic decision process rather than a simple transformation to the accuracy scale. In light of these changes, we present the full mathematical details of the model, indicating the nature of these changes. It is important to note that the modified model reduces to the original model as applied in previous applications. Previous applications used connection strengths set to the default of 1.0 (in which case the modified synaptic depression equation reduces to the original equation). Also, previous applications only examined a single class of stimulus (in which case the stochastic decision process produces the same accuracy predictions as the original decision equation). Thus, prior modeling results would be replicated with the modified model.

### *A.1. Network architecture*

As seen in Fig. 2, the three layers represent visual (location specific), orthographic (regardless of location), and lexical-semantic (regardless of location) information. Besides the feed-forward connections, there are feedback connections from lexical-semantics to orthography, giving the model attractor dynamics. All of the connections between layers are excitatory. In contrast, connections within layers are inhibitory and fixed to be all-to-all, reflecting inhibitory interneurons (O'Reilly and Munakata, 2000), producing masking effects at all layers. Inhibitory connections do not bridge across different locations at the bottom layer, capturing retinotopic masking, in contrast to meta-masking, which occurs through inhibition at higher layers. As applied to repetition priming, there is a node representing each word at each layer, and a node representing the mask at the visual layer (a pattern mask has no orthographic or lexical-semantic content). For the visual layer there are separate nodes for each word at each screen location where that word might appear. For associative priming, the association between two items is captured by a shared lexical-semantic representation, which is a simplification of a fully distributed representation with shared semantic features (Masson, 1995).

Connections weights between nodes (the $w$ parameters in Eq. (A1)) are assumed to exist, with value 1.0, or not exist (value 0), as dictated by the condition being simulated. The feedback from the lexical-semantic layer onto the orthographic layer is scaled by the $F$ parameter in Eq, (A1), set to a default value of 0.25 (the orthographic layer nodes do not feedback onto visual layer nodes). For the visual layer, 5 nodes were simulated (the prime, the target, the mask, the left choice, and the right choice). The prime, target, and mask were all located in the center of the screen and for this reason they inhibited/masked each other. In contrast, the left and right choices appeared in different screen positions than the other stimuli and had no competitors aside from self-inhibition. Experiments 1 and 2 presented two copies of the prime word and rather than explicitly simulating two identically behaving visual nodes for each prime, the output of the single prime node was multiplied by 2, achieving the same mathematical result. It is important to note that this does not produce a doubling in the orthographic activation of the prime word owing to a non-linear activation function. For the orthographic and lexical-semantic layers, up to 6 nodes per layer were simulated (i.e., representations unique to the prime, shared between prime and target, shared between prime and foil, unique to the target, unique to the foil, and the closed class word used in Experiment 3). However, a particular condition only required a subset of these nodes. For instance, for repetition priming with the target primed condition, 2 orthographic nodes were needed and 2 lexical-semantic nodes were needed (at each layer one node for a representation shared between prime and target and one for a representation unique to the foil). This can be contrasted with associative priming with the target primed condition, which required 3 orthographic nodes (the prime's orthography, the target's orthography, and the foil's orthography) and 4 lexical-semantic nodes (meaning unique to the prime, meaning shared between prime and target, meaning unique to the target, and the meaning of the foil).

### *A.2. Network dynamics*

Each node approximates an assembly of neurons having similar inputs and outputs. The size of an assembly is captured by the representational magnitude, $m$, in Eq. (A1), representing the relative number of neurons (i.e., relative to some typical situation) that are similarly connected and thus respond in a similar manner. By default, the $m$ parameters are set to 1.0, representing the number of neurons represented by each node for the typical stimulus (e.g., high frequency English words) at each layer of the model. The reported experiments modeled associative/semantic priming by breaking the $m = 1$ lexical-semantic neurons that represent a word into two different pools of neurons to differentiate between meaning that is shared between the prime and a primed choice word ($m = m_{shared}$) versus meaning that is unique to the prime or primed choice word ($m = 1 - m_{shared}$).

$$\frac{\Delta v_i^n}{S_n} = (1 - v_i^n)\left\{\sum_{\forall j} m_j w_{ij} o_j^{n-1} + F\sum_{\forall k} m_k w_{ik} o_k^{n+1}\right\} - v_i^n\left\{L + I\sum_{\forall l} m_l o_l^n\right\} \tag{A1}$$

Eq. (A1) specifies the dynamic update at each millisecond for the membrane potential of node $i$ at layer $n$ according to excitatory inputs, found in the first set of curly brackets, which drive membrane potential toward the maximum of 1.0, versus inhibitory inputs, found in the second set of curly brackets, which drive membrane potential toward the minimum of 0.0. The parameter $S$ captures the speed at which layer $n$ updates to reflect the current set of inputs versus maintaining the membrane potential of the previous millisecond. The excitatory inputs are further broken down into bottom-up inputs, found in the summation over all nodes $j$ in layer $n$-1, with connection weights $w_{ij}$, output activity $o_j$, and representational magnitude $m_j$, versus top-down inputs, found in the summation over all nodes $k$ in layer $n + 1$, with connection weights $w_{ik}$, output activity $o_k$, and representational magnitude $m_k$. The inhibitory inputs consist of a constant leak current $L$ and all-to-all within-layer mutual inhibition, found in the summation over all nodes $l$ within layer $n$ (with the exception that inhibition depends on screen position for the visual layer). The terms in this summation are over the output $o_l$ of all nodes, with each scaled by the appropriate representational magnitude $m_l$.

Eq. (A2) captures the threshold nature of neurons, with the output, $o$, of neurotransmitter to each synapse set to 0 when membrane potential is below the firing threshold $\theta$, but proportional to the average firing rate, $v - \theta$, when membrane potential is above threshold.

$$o = \begin{cases} (v - \theta)a & v > \theta \\ 0 & v \leqslant \theta \end{cases} \tag{A2}$$

Output activity is scaled by the variable $a$, which quantifies the available presynaptic resources, capturing neural habituation through synaptic depression. Evidence suggests that synaptic depression is primarily a pre-synaptic phenomenon (Simons-Weidenmaier, Weber, Plappert, D Pilz, & Schmid, 2006) and, furthermore, that the dynamics of synaptic depression are cell-wide, rather than unique to each synapse (Armbruster & Ryan, 2011). Thus, each cell has a limited amount of neurotransmitter available for signaling receiving cells and the $a$ variable represents the current level of available neurotransmitter. Dynamic changes in the level of available transmitter are specified by Eq. (A3).

$$\frac{\Delta a_i}{S_n} = R(1 - a_i) - Do_i \sum_{\forall j} w_{ji} \tag{A3}$$

Eq. (A3) stipulates the change each millisecond in the level of available synaptic resources for cell $i$. This change depends on a recovery parameter, $R$, indicating how rapidly resources recovery to the maximum of 1.0, and a depletion parameter, $D$, indicating how rapidly resources are depleted as a function of the current output activity. **CHANGE 1:** Eq. (A3) is slightly changed from the equation reported in Huber and O'Reilly (2003) by multiplying the current output, $o_i$, by the summation of all receiving weights, $w_{ji}$, of all the nodes $j$ that cell $i$ projects to. This change reflects evidence that synaptic plasticity results in pre-synaptic mechanisms that modify the probability of neurotransmitter release (Zakharenko et al., 2003). Thus, a cell that is more strongly connected to receiving cells is one that more rapidly releases its neurotransmitter and more rapidly experiences cell-wide depletion of neurotransmitter. For this reason, the effective depletion constant for Eq. (A3) is the parameter depletion parameter $D$, scaled by the level of connectivity for the cell, as reflected in the summation term. This modification has important consequences for the model and contributes to the prediction that habituation will occur more rapidly for well-learned stimuli (i.e., stimuli with neural representations that are more strongly connected). However, it is important to note previously reported simulations with nROUSE are entirely consistent with this change because this summation term would have been 1.0 for those simulations (i.e., prior nROUSE simulations kept the connection weights fixed at 1.0).

### *A.3. Decision process*

A race process is assumed to underlie choice behavior. Specifically, the choice word identified more quickly is selected. To implement this decision rule, the identification latency for each word is the time in milliseconds when lexical-semantic output achieves a maximum value before beginning to decrease owing to habituation (in cases where the lexical-semantic representation of a word is broken into different parts, such as with associative priming, these parts are summed before determining whether the lexical-semantic output has achieved a maximum). The dynamics of nROUSE are fully deterministic and assumptions of variability are needed to produce accuracy other than 0% or 100%. **CHANGE 2:** The original model transformed the difference in identification latencies for the target and foil choice alternatives into the accuracy scale using the logistic equation. However, this solution does not specify the distributional properties of identification latency. Instead, the current model assumes that identification latency is normally distributed with a mean equal to the simulation-determined identification latency, $\mu_t$, and variance that is exponentially proportional to the mean latency as indicated by Eq. (A4) with the noise parameter $\xi$.

$$\sigma_t^2 = e^{\xi \mu_t} \tag{A4}$$

$$\text{proportion correct} = \int_0^\infty \frac{1}{(\sigma_T^2 + \sigma_F^2)^{1/2}\sqrt{2\pi}} e^{-\frac{(x+\mu_T-\mu_F)^2}{2(\sigma_T^2+\sigma_F^2)}} dx \tag{A5}$$

If modeling reaction time distributions, the equations for an independent race process involve multiplying the density function for the winning racer by one minus the cumulative distribution for the losing racer. However, if one only cares about average accuracy from the race process, regardless of finishing time, then a simpler expression of accuracy can be derived by considering the difference distribution between the racers. More specifically, if the difference distribution is foil identification latency minus target identification latency, accuracy is equal to the integral from a difference of 0 to infinity (i.e., one minus the cumulative difference distribution) because any positive values of this difference indicate that the target won the race. Based on the assumption of normally distributed identification latencies, Eq. (A5) is the cumulative normal distribution for the difference between the foil and target latencies that calculates accuracy. Combining these two equations, accuracy is calculated from the noise constant $\xi$, the time to identify the target $\mu_T$, and the time to identify the foil $\mu_F$.

For a given stimulus class, priming manipulations primarily affect the mean difference between the identification latency of the target versus the foil (i.e., the primed choice is made to be faster or slower, depending on prime duration, but these priming manipulations do not affect the identification latency of the unprimed choice). Therefore, the predicted effects of priming manipulations for a given class of stimulus are the same regardless of the assumed relationship between the means and variances of the identification latencies. However, with connection strengths less than 1.0, such as with non-words, identification latency to both primed and unprimed choices may slow down considerably (e.g., 300 ms rather than 200 ms). Because Experiment 1 directly compared conditions with connection weights of 1.0 versus ones with smaller weights, the assumption of an exponential relationship between the mean and variance dictates that performance should be worse for stimuli that are in general identified more slowly (in contrast, the original logistic equation incorrectly produces higher accuracy for non-words, owing to a larger mean difference between the identification latency of the target versus the foil).

### *A.4. Applying nROUSE to Experiment 1–3*

Many of nROUSE's parameters relate to the basic operations of neurons and these parameters should not vary across different applications of the model. These previously published default parameters were: $L = 0.15$ (leak current), $D = 0.324$ (rate of synaptic depletion), $R = 0.022$ (rate of synaptic recovery), $\theta = 0.15$ (firing threshold), and $F = 0.25$ (feedback strength). In addition, the integration time constant for the top layer was set to a default value of $S_3 = 0.015$ (speed of lexical-semantic integration), which is a value proven to work for a variety of circumstances (in general, this value needs to be smaller than the time constant of the lower layers of the model, but otherwise does not greatly affect the behavior of the model).

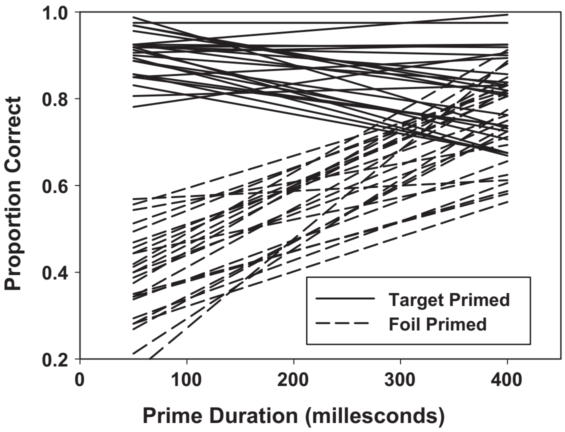
The remaining parameters were allowed to freely vary to capture the results of Experiment 1 in an attempt to describe repetition priming for the visual characteristics used in the current study (e.g., font size and type, visual angle, visual contrast, etc.). The seven free parameters in applying the model to Experiment 1 included the visual ($S_1 = 0.0294$) and orthographic layer ($S_2 = 0.0609$) time constants, the inhibition parameter ($I = 0.9844$) and the noise constant ($\xi = 0.0302$). Non-words were assumed to differ from words only in the connection strength between the orthographic representation of the non-word and the lexical-semantic representation of any orthographically similar valid words (connection strength of $w_{32} = w_{23} = 0.9339$ for pronounceable non-words and $w_{32} = w_{23} = 0.8710$ for non-pronounceable non-words). Inverted words were assumed to differ from words only in the connection strength between the visual representation of the inverted word and the correct orthographic representation of the inverted word (connection strength of $w_{21} = 0.6195$). This describes MODEL 1, whereas the $w$ parameters were all fixed at 1.0 for MODEL 2. Instead, of using connection strength, MODEL 2 captured accuracy differences for the four different kinds of stimuli by using four different noise constants.

Two free parameters were needed to capture the manner in which Experiment 2 differed from Experiment 1. Experiment 2 was identical to Experiment 1, and so the time constants and inhibition parameter were kept at the values that best-fit the results of Experiment 1. However, different groups of participants may vary in overall accuracy, and the noise parameter ($\xi = 0.0304$ for Experiment 2) was allowed to freely vary to capture these differences. To capture the extent of associative priming, the representational magnitude of the shared lexical-semantic content ($m_{shared} = 0.1687$) was free to vary, but applied equally to forward and backward associative priming. This describes MODEL 1, whereas MODEL 2 assumed that a prime word directly activates the layer 2 representation of the target with a connection strength less than 1.0 ($w_{12} = 0.3019$) rather than assuming a shared layer 3 representation. This is shown in Fig. 4 with the dashed connection, although for reasons of space, MODEL 2 in Fig. 4 does not show that the prime was also connected to its own layer 2 node with the usual default strength of 1.0 and that the target was also connected to the prime's layer 2 node with the same $w_{12}$ as used for the prime to target relation. In theory, the two different layer 1 to layer 2 connections for each word in MODEL 2 should have been summed, dictating that the word's visual response habituate more quickly as specified in Eq. (A3). Nevertheless, to keep MODEL 2 consistent with the repetition priming results of Experiment 1, it was assumed that visual layer habituation occurred according to a summed connection strength of 1.0 rather than $1 + w_{12}$ (doing otherwise makes it more difficult for MODEL 2 to fit the data).

Three free parameters were needed to capture the manner in with Experiment 3 differed from Experiment 1. In modeling Experiment 3, the noise constant was again a free parameter to capture average accuracy for the subjects who participated in Experiment 3 ($\xi = 0.0288$). Keeping all other parameters fixed at the same values used to explain the word priming data of Experiment 1, only two additional parameters were allowed to vary. Because Experiment 3 used a lower case prime presented at the midline, rather than upper case primes presented above and below the midline, the connection strength between the visual response to the prime and corresponding orthography was allowed to freely vary ($w_{21} = 1.956$, which can be compared to the default value of 1.0 assumed for Experiment 1 and 2). The third free parameter was the representational magnitude of the shared lexical-semantic content ($m_{shared}$) which was 0.1530, serving the same role as for MODEL 1 of Experiment 2. This described MODEL 1, whereas MODEL 2 dropped the assumption of a shared layer 3 representation for the expected conditions, instead modeling the role of expectations through direct connections between the prime's visual response and the target's layer 2 response in the same manner as MODEL 2 as applied to Experiment 2.

## Appendix B. Individual differences

To explore the role of individual differences, we replicated a subset of the word repetition priming conditions from Experiment 1 with 100 trials per condition for each participant. This experiment examined primes presented for 50 ms or 400 ms, either with target priming or foil priming. Additional details of this experiment are reported by (Potter et al., submitted for publication). As seen in the figure below, all of the 25 participants exhibited a large positive priming effect (target primed > - foil primed) following 50 ms primes (the raw data from the target primed and foil primed conditions are completely non-overlapping). In moving from the 50 ms prime duration conditions to the 400 ms prime duration conditions, the difference between the target primed and foil primed conditions decreased for all 25 participants.



## References

Anderson, J. R. (1991). The adaptive nature of human categorization. *Psychological Review, 98*, 409–429.

Armbruster, M., & Ryan, T. A. (2011). Synaptic vesicle retrieval time is a cell-wide rather than individual-synapse property. *Nature Neuroscience, 14*, 824–826.

Bodner, G. E., & Masson, M. E. J. (2003). Beyond spreading activation: An influence of relatedness proportion on masked semantic priming. *Psychonomic Bulletin & Review, 10*, 645–652. http://dx.doi.org/10.3758/Bf03196527.

Brants, T., & Franz, A. (2006). *Web 1T 5-gram Version 1*.

Burt, J. S., Kipps, T. J., & Matthews, J. R. (2014). Repetition in visual word identification: Benefits and costs. *Quarterly Journal of Experimental Psychology, 67*, 1986–2009. http://dx.doi.org/10.1080/17470218.2014.896386.

Carr, T. H., & Dagenbach, D. (1990). Semantic priming and repetition priming from masked words - Evidence for a center-surround attentional mechanism in perceptual recognition. *Journal of Experimental Psychology-Learning Memory and Cognition, 16*, 341–350.

Coltheart, M. (1981). The MRC psycholinguistic database. *Quarterly Journal of Experimental Psychology Section a-Human Experimental Psychology, 33*, 497–505.

Dagenbach, D., Carr, T. H., & Wilhelmsen, A. L. (1989). Task-induced strategies and near-threshold priming - Conscious influences on unconscious perception. *Journal of Memory and Language, 28*, 412–443.

Davelaar, E. J., Tian, X., Weidemann, C. T., & Huber, D. E. (2011). A habituation account of change detection in same/different judgments. *Cognitive Affective & Behavioral Neuroscience, 11*, 608–626. http://dx.doi.org/10.3758/s13415-011-0056-8.

de Wit, B., & Kinoshita, S. (2015a). The masked semantic priming effect is task dependent: Reconsidering the automatic spreading activation process. *Journal of Experimental Psychology-Learning Memory and Cognition, 41*, 1062–1075.

de Wit, B., & Kinoshita, S. (2015b). An RT distribution analysis of relatedness proportion effects in lexical decision and semantic categorization reveals different mechanisms. *Memory & Cognition, 43*, 99–110.

Dukewich, K. R. (2009). Reconceptualizing inhibition of return as habituation of the orienting response. *Psychonomic Bulletin & Review, 16*, 238–251.

Duncan, J., Ward, R., & Shapiro, K. (1994). Direct measurement of attentional dwell time in human vision. *Nature, 369*, 313–315.

Evett, L. J., & Humphreys, G. W. (1981). The use of abstract graphemic information in lexical access. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 33A*, 325–350. http://dx.doi.org/10.1080/14640748108400797.

Fischler, I. (1977). Semantic facilitation without association in a lexical decision task. *Memory & Cognition, 5*, 335–339.

Gomez, P., Perea, M., & Ratcliff, R. (2013). A diffusion model account of masked versus unmasked priming: Are they qualitatively different? *Journal of Experimental Psychology-Human Perception and Performance, 39*, 1731–1740.

Grainger, J., & Jacobs, A. M. (1999). Temporal integration of information in orthographic priming. *Visual Cognition, 6*, 461–492.

Griffiths, T. L., Steyvers, M., & Tenenbaum, J. B. (2007). Topics in semantic representation. *Psychological Review, 114*, 211–244.

Hochhaus, L., & Johnston, J. C. (1996). Perceptual repetition blindness effects. *Journal of Experimental Psychology-Human Perception and Performance, 22*, 355–366. http://dx.doi.org/10.1037/0096-1523.22.2.355.

Hogben, J. H., & Di Lollo, V. (1974). Perceptual integration and perceptual segregation of brief visual stimuli. *Vision Research, 14*, 1059–1069. http://dx.doi.org/10.1016/0042-6989(74)90202-8.

Hubel, D. H. (1988). *Eye, brain, and vision*. New York: Scientific American Library. Distributed by W.H. Freeman.

Huber, D. E. (2008). Immediate priming and cognitive aftereffects. *Journal of Experimental Psychology-General, 137*, 324–347. http://dx.doi.org/10.1037/0096-3445.137.2.324.

Huber, D. E. (2015). Using continual flash suppression to investigate cognitive aftereffects. *Consciousness and Cognition, 35*, 30–32. http://dx.doi.org/10.1016/j.concog.2015.04.017.

Huber, D. E., Clark, T. F., Curran, T., & Winkielman, P. (2008). Effects of repetition priming on recognition memory: Testing a perceptual fluency-disfluency model. *Journal of Experimental Psychology-Learning Memory and Cognition, 34*, 1305–1324. http://dx.doi.org/10.1037/a0013370.

Huber, D. E., & Cousineau, D. (2004). A race model of perceptual forced choice reaction time. In *Proceedings of the 25th annual conference of the cognitive science society* (pp. 593–598). Hillsdale, NJ: Erlbaum Associates.

Huber, D. E., & O'Reilly, R. C. (2003). Persistence and accommodation in short-term priming and other perceptual paradigms: Temporal segregation through synaptic depression. *Cognitive Science, 27*, 403–430. http://dx.doi.org/10.1207/s15516709cog2703_4.

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

Huber, D. E., Tian, X., Curran, T., O'Reilly, R. C., & Woroch, B. (2008). The dynamics of integration and separation: ERP, MEG, and neural network studies of immediate repetition effects. *Journal of Experimental Psychology-Human Perception and Performance, 34*, 1389–1416. http://dx.doi.org/10.1037/a0013625.

Humphreys, G. W., Besner, D., & Quinlan, P. T. (1988). Event perception and the word repetition effect. *Journal of Experimental Psychology: General, 117*, 51–67. http://dx.doi.org/10.1037/0096-3445.117.1.51.

Irwin, K. R., Huber, D. E., & Winkielman, P. (2010). Automatic affective dynamics: an activation-habituation model of affective assimilation and contrast. *Modeling Machine Emotions for Realizing Intelligence, 1*, 17–34. http://dx.doi.org/10.1007/978-3-642-12604-8.

Johnston, W. A., & Hawley, K. J. (1994). Perceptual Inhibition of Expected Inputs - the Key That Opens Closed Minds. *Psychonomic Bulletin & Review, 1*, 56–72.

Jones, M. N., & Mewhort, D. J. K. (2007). Representing word meaning and order information in a composite holographic lexicon. *Psychological Review, 114*, 1–37.

Kahan, T. A. (2000). Negative priming from masked words: Retrospective prime clarification or center-surround inhibition? *Journal of Experimental Psychology-Learning Memory and Cognition, 26*, 1392–1410.

Kanwisher, N. G. (1987). Repetition blindness - Type recognition without token Individuation. *Cognition, 27*, 117–143. http://dx.doi.org/10.1016/0010-0277(87)90016-3.

Kucera, H., & Francis, W. N. (1967). *Computational analysis of present day American English*. Providence, RI: Brown University Press.

Landauer, T. K., & Dumais, S. T. (1997). A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychological Review, 104*, 211–240.

Lucas, M. (2000). Semantic priming without association: A meta-analytic review. *Psychonomic Bulletin & Review, 7*, 618–630.

Lund, K., & Burgess, C. (1996). Producing high-dimensional semantic spaces from lexical co-occurrence. *Behavior Research Methods Instruments & Computers, 28*, 203–208.

Marr, D. (1982). *Vision: A computational investigation into the human representation and processing of visual information*. San Francisco: W.H. Freeman.

Masson, M. E. J. (1995). A distributed-memory model of semantic priming. *Journal of Experimental Psychology-Learning Memory and Cognition, 21*, 3–23.

Masson, M. E. J. (2004). When words collide: Facilitation and interference in the report of repeated words from rapidly presented lists. *Journal of Experimental Psychology-Learning Memory and Cognition, 30*, 1279–1289.

McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception. 1. An account of basic findings. *Psychological Review, 88*, 375–407.

Meyer, D. E., & Schvaneveldt, R. W. (1971). Facilitation in recognizing pairs of words: Evidence of a dependence between retrieval operations. *Journal of Experimental Psychology, 90*, 227–234.

Meyer, D. E., Schvaneveldt, R. W., & Ruddy, M. G. (1974). Functions of graphemic and phonemic codes in visual word-recognition. *Memory & Cognition, 2*, 309–321.

Nelson, D. L., McEvoy, C. L., & Schreiber, T. A. (1998). The University of South Florida word association, rhyme and word fragment norms.

New, B., & Grainger, J. (2011). On letter frequency effects. *Acta Psychologica, 138*, 322–328.

Norris, D., & Kinoshita, S. (2008). Perception as evidence accumulation and Bayesian inference: Insights from masked priming. *Journal of Experimental Psychology-General, 137*, 434–455.

O'Reilly, R. C., & Munakata, Y. (2000). *Computational explorations in cognitive neuroscience: Understanding the mind by simulating the brain*. Cambridge, Mass.: MIT Press.

Pecher, D., Zeelenberg, R., & Raaijmakers, J. G. W. (1998). Does pizza prime coin? Perceptual priming in lexical decision and pronunciation. *Journal of Memory and Language, 38*, 401–418.

Pecher, D., Zeelenberg, R., & Raaijmakers, J. G. W. (2002). Associative priming in a masked perceptual identification task: Evidence for automatic processes. *Quarterly Journal of Experimental Psychology Section A-Human Experimental Psychology, 55*, 1157–1173.

Posner, M. I., & Cohen, Y. (1984). Components of visual orienting. In H. Bouma & D. Bouwhuis (Eds.). *Attention and performance* (Vol. X, pp. 531–556). Erlbaum.

Potter, K. W., Donkin, C., & Huber, D. E. (tter et al., submitted for publication). Testing a perceptual fluency/disfluency model of priming with a model of response time and choice. *Cognitive Psychology*. Submitted for publication.

Potter, M. C., & Levy, E. I. (1969). Recognition memory for a rapid sequence of pictures. *Journal of Experimental Psychology, 81*, 10.

Ratcliff, R., & McKoon, G. (1997). A counter model for implicit priming in perceptual word identification. *Psychological Review, 104*, 319–343.

Ratcliff, R., McKoon, G., & Verwoerd, M. (1989). A bias interpretation of facilitation in perceptual identification. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 15*, 378–387.

Raymond, J. E., Shapiro, K. L., & Arnell, K. M. (1992). Temporary suppression of visual processing in an RSVP task - An attentional blink. *Journal of Experimental Psychology-Human Perception and Performance, 18*, 849–860. http://dx.doi.org/10.1037/0096-1523.18.3.849.

Riefer, D. M., & Batchelder, W. H. (1988). Multinomial modeling and the measurement of cognitive-processes. *Psychological Review, 95*, 318–339.

Rieth, C. A., & Huber, D. E. (2010). Priming and habituation for faces: Individual differences and inversion effects. *Journal of Experimental Psychology-Human Perception and Performance, 36*, 596–618. http://dx.doi.org/10.1037/a0018737.

Rieth, C. A., & Huber, D. E. (2013). Implicit learning of spatiotemporal contingencies in spatial cueing. *Journal of Experimental Psychology-Human Perception and Performance, 39*, 1165–1180.

Rusconi, P., & Huber, D. E. (sconi and Huber, submitted for publication). The perceptual wink model of non-switching attentional blink tasks. *Psychonomic Bulletin & Review*. Submitted for publication.

Simons-Weidenmaier, N. S., Weber, M., Plappert, C. F., D Pilz, P. K., & Schmid, S. (2006). Synaptic depression and short-term habituation are located in the sensory part of the mammalian startle pathway. *BMC Neuroscience, 7*(38). http://dx.doi.org/10.1186/1471-2202-7-38. PMCID: PMC1479352.

Thomson, A. M., & West, D. C. (1993). Fluctuations in pyramid excitatory postsynaptic potentials modified by presynaptic firing pattern and postsynaptic membrane-potential using paired intracellular-recordings in rat neocortex. *Neuroscience, 54*, 329–346.

Tian, X., & Huber, D. E. (2010). Testing an associative account of semantic satiation. *Cognitive Psychology, 60*, 267–290. http://dx.doi.org/10.1016/j.cogpsych.2010.01.003.

Tian, X., & Huber, D. E. (2013). Playing "Duck Duck Goose" with neurons: Change detection through connectivity reduction. *Psychological Science, 24*, 819–827. http://dx.doi.org/10.1177/0956797612459765.

Tsodyks, M. V., & Markram, H. (1997). The neural code between neocortical pyramidal neurons depends on neurotransmitter release probability. *Proceedings of the National Academy of Sciences USA, 94*, 719–723. http://dx.doi.org/10.1073/pnas.94.2.719.

Wagenmakers, E. J. M., Zeelenberg, R., Steyvers, M., Shiffrin, R., & Raaijmakers, J. G. W. (2004). Nonword repetition in lexical decision: Support for two opposing processes. *Quarterly Journal of Experimental Psychology Section a-Human Experimental Psychology, 57*, 1191–1210.

Weidemann, C. T., Huber, D. E., & Shiffrin, R. M. (2008). Prime diagnosticity in short-term repetition priming: Is primed evidence discounted, even when it reliably indicates the correct answer? *Journal of Experimental Psychology-Learning Memory and Cognition, 34*, 257–281.

Yin, R. K. (1969). Looking at upside-down Faces. *Journal of Experimental Psychology, 81*, 141.

Yue, X. M., Tjan, B. S., & Biederman, I. (2006). What makes faces special? *Vision Research, 46*, 3802–3811.

Zakharenko, S. S., Patterson, S. L., Dragatsis, I., Zeitlin, S. O., Siegelbaum, S. A., Kandel, E. R., & Morozov, A. (2003). Presynaptic BDNF required for a presynaptic but not postsynaptic component of LTP at hippocampal CA1-CA3 synapses. *Neuron, 39*, 975–990.