Nick Blauch
Independent Study in the Cognitive Experiments, Models, and Neuroscience Laboratory
Advisor: David Huber
Fall 2015

Stochastic nROUSE: Investigations into Sources of Perceptual Variability

The nROUSE model has successfully accounted for previously conflicting results in repetition priming paradigms. Utilizing the dynamics of synaptic depression within a simple neural network model, nROUSE models behavioral findings and provides an intuitive understanding of repetition priming effects. However, in a perceptual horse race between target and foil alternatives, the nROUSE model produces deterministic racer latencies. This is a twofold problem. First, in order to produce non-extreme accuracy, it requires that the nROUSE model incorporate variability in detecting which racer (i.e, target or foil) reaches peak activation first. Second, it means that the nROUSE model is unfit to account for variable reaction times out of the box. I explore sources of perceptual variability in order to generate stochastic nROUSE racer latencies. In doing so, I explore the possibilities of 1) overcoming the need for a stochastic race-detection process and 2) modeling reaction times in repetition priming experiments.

1. Introduction

Repetition priming is typically studied in rapid-serial-visual-presentation (RSVP) tasks. In these tasks, a stream of stimuli are shown in rapid succession, usually for very brief periods of time. In the prototypical repetition priming experiment, a prime is first displayed which is either a target or foil word, followed by presentation of the target word, followed by a mask, followed by the presentation of the two choices of a 2-alternative forced choice decision. Basic intuition suggests that the presentation of a target prime should render identification of the target word easier than the unprimed identification task. However, in measuring performance across a range of prime durations, a more complicated pattern emerges. As shown in Figure 1, for short prime durations, as expected, the presentation of a target prime aids performance above baseline, and the presentation of a foil prime hinders performance below baseline. However, for longer prime durations the effect reverses; a target prime hinders performance, and a foil prime aids performance.
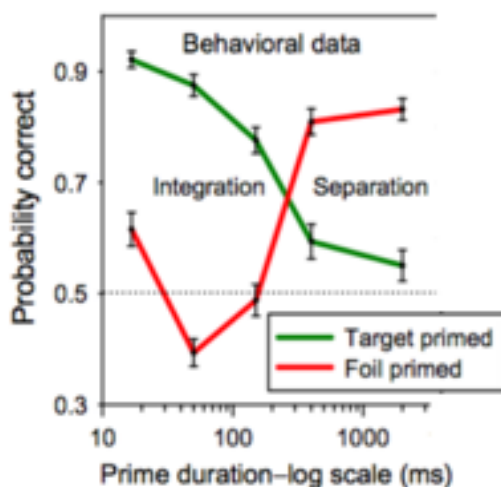
Figure 1. Repetition priming. Probability correct is plotted as a function of the prime duration for both target and foil primes. The plot shows the characteristic crossover pattern found in repetition priming experiments which manipulate prime duration. Adopted from Huber 2014.

This pattern of performance as a function of prime duration was first formally explained by Huber's ROUSE model. This model, called Responding Optimally with Unknown Sources of Evidence, used a bayesian framework to describe why and in what manner a perceptual system would discount prior evidence (i.e, the perceptual input of a prime) to support temporal separation of different stimuli. Huber later reformulated the model in a neural network, incorporating the dynamics of synaptic depression to naturally account for priming data; this model was coined nROUSE, a Neural mechanism for Responding Optimally with Unknown Sources of Evidence. This model is the topic of the discussion to follow.

The nROUSE model is a three-layer neural network consisting of visual, orthographic, and lexical-semantic processing layers. There exists one visual node per word for each location in which that word appears in a given experiment, and one node per word at each of the orthographic and lexical-semantic layers. Connections consist of inhibitory connections within layers and feedforward and feedback excitatory connections between layers. The architecture is displayed in Figure 2.
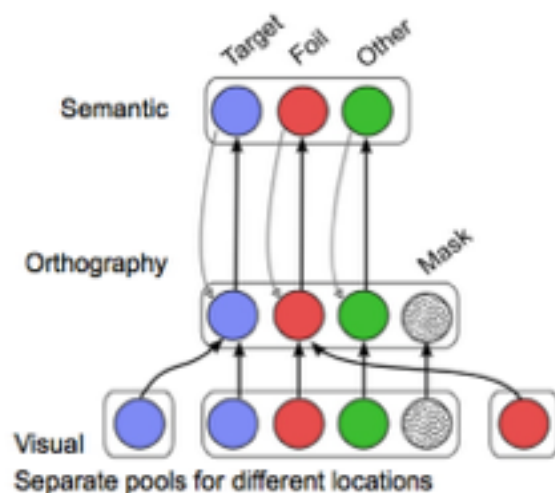

Figure 2. Architecture of the neural network of the nROUSE model. Adopted from Huber 2014.

The nROUSE model uses a horse-race peak-activation decision rule to predict what word is chosen under a given set of experimental conditions. What this means is that the activity of the highest (lexical-semantic) layer of processing is tracked over time, and the first node to reach peak activation is said to be the winner of the horse-race. Throughout this text, I refer to the time it takes a given lexical-semantic node (i.e, target or foil) to reach peak activation as the (target or foil) latency. But there is a catch. The nROUSE model produces deterministic latencies. So, under the same set of experimental conditions and modeling parameters, the model will always output the same target and foil latencies, which means that accuracy will always be 100% or 0%, depending on whether the target or foil latency is shorter. Thus, to model human performance — which rarely averages to 0% or 100% over many trials — the model assumes variability in the perceptual system's ability to detect the shorter latency. A normal distribution of

(foil-target) latency difference is assumed with mean μ given by the output foil - target latency, and variance σ^2 given by:

$$\sigma^2 = \sum_i e^{N \cdot l_i}$$

(1)

Where summation occurs over the vector of target and foil latencies l, and N is a fitted noise parameter. The proportion correct $p(c)$ is then given by 1 — the cumulative density of the distribution of (foil — target latency), evaluated at 0, where the latencies would be equal. This is an effective method for generating modeled accuracies which map onto behavioral findings; however, it requires an additional noise parameter whose ground truth is unspecified. I explore alternative methods for generating reasonable accuracies from the nROUSE model; these methods assume no variability in detecting the winner of the horse-race, such that the word with the shorter latency is always chosen. It is thus the goal of these methods to generate variable latencies.

This goal has a second application. Assuming that the shorter latency is always chosen allows for a simple interpretation of reaction time within the context of the nROUSE model. Many cognitive models assume that reaction time can be broken down into two distinct non-overlapping components: the evidence accumulation or decision component, and the motor-delay. It is easy to see that, within this framework, the shorter latency produced by nROUSE should correspond to the decision component. For the sake of this work, I assume nothing about the motor-delay component of reaction time. My goal is to produce patterns of variability within the decision component which map onto patterns of variability within behavioral reaction times.

In Huber & O'Reilly (2003), which introduced the nROUSE model, the authors listed several possible sources of variability which could be exploited in the nROUSE model in future work. These sources are: selective attention, processing noise, criterial shifts, encoding variability, and item differences. In this paper, I explore item differences and encoding variability through stochastic simulations of the nROUSE model. For each simulation, I investigate 1) whether the noise can overcome the need for a stochastic race-detection process, and 2) whether the noise might aid in modeling reaction times.

2. Item differences

The idea behind item differences as a source of variability is that different words may have different connection weights dependent on the extent of one's experience with each word. In accordance with this notion, Wagenmakers et al (2000) found that identification accuracy was higher for words with a higher written language frequency. When words are chosen for repetition-priming experiments, target and foil words are sometimes matched by their written language frequency, in order to avoid possible item differences. However, even when words are matched in this way, item differences likely exist within smaller populations or at the individual level. Thus, I first explore item differences through simulations in which the orthographic to lexical-semantic connections are drawn from a normal distribution with $\mu = 1$ and $\sigma = .3$. The standard deviation value was chosen arbitrarily as a value which would allow for large item differences to be examined. Feedback connections were set to the same value as feedforward connections, to maintain symmetry. Figure 3 shows the results of item differences as realized through variation in orthographic to lexical semantic connection weights.
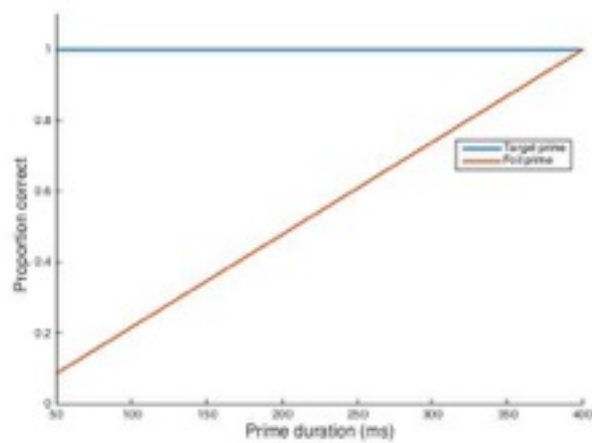
a.
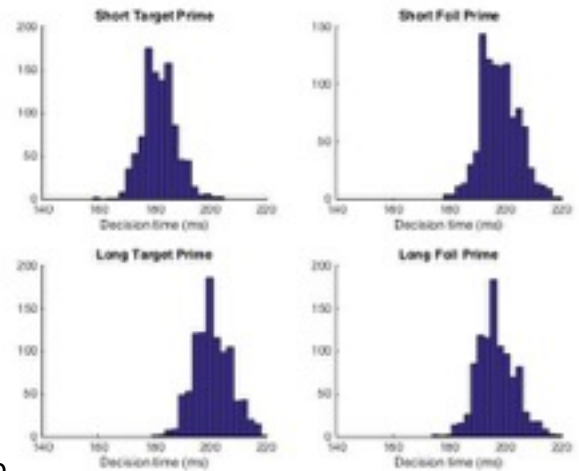
b.

Figure 3. Item differences realized through stochastic orthographic to lexical-semantic connection strength. Connection weights are normally distributed with $\mu = 1$ and $\sigma = .3$.

Proportion correct vs. prime duration is shown in (a). 20-bin histograms of decision time (as defined by the shortest latency) for 4 conditions are shown in (b). Short primes are 50 ms, and long primes are 400 ms.

From 3a, it is clear that this form of item differences is not sufficient to produce variable latencies which result in the characteristic priming "crossover" pattern, as seen in Figure 1. From 3b, it is clear that this form of item differences does not produce a right-skew in decision times, as found in behavioral reaction times.

I next examine item differences in visual to orthographic connection strength. Again, connections are drawn from a normal distribution with $\mu = 1$ and $\sigma = .3$, and symmetric feedforward and feedback connections are used.
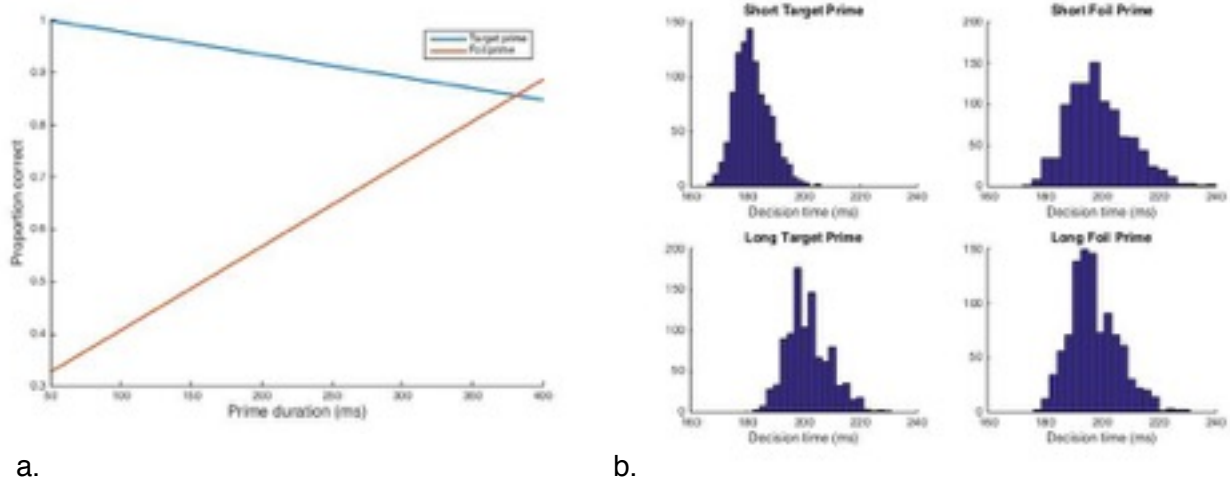
a.

b.

Figure 4. Item differences realized through stochastic visual to orthographic connection strength.Connection weights are normally distributed with $\mu = 1$ and $\sigma = .3$. Proportion correct vs. prime duration is shown in (a). 20-bin histograms of decision time (as defined by the shortest latency) for 4 conditions are shown in (b). Short primes are 50 ms, and long primes are 400 ms.

From 4a, we see that this form of item differences does produce the characteristic priming "crossover" pattern, as seen in Figure 1. From 4b, however, we see no evidence of a right-skew in decision times, as found in behavioral reaction times.

3. Visual encoding variability

The next stochastic process I examined is visual encoding variability: the notion that the visual system may not fully encode the stimulus, especially for short durations. I first examine random encoding variability, where visual input strength is determined individually at each visual presentation, excluding the mask (i.e, prime, target, choices). Visual input is drawn from a normal distribution with $\mu = 1$ and $\sigma = .3$. Results are shown in Figure 5.
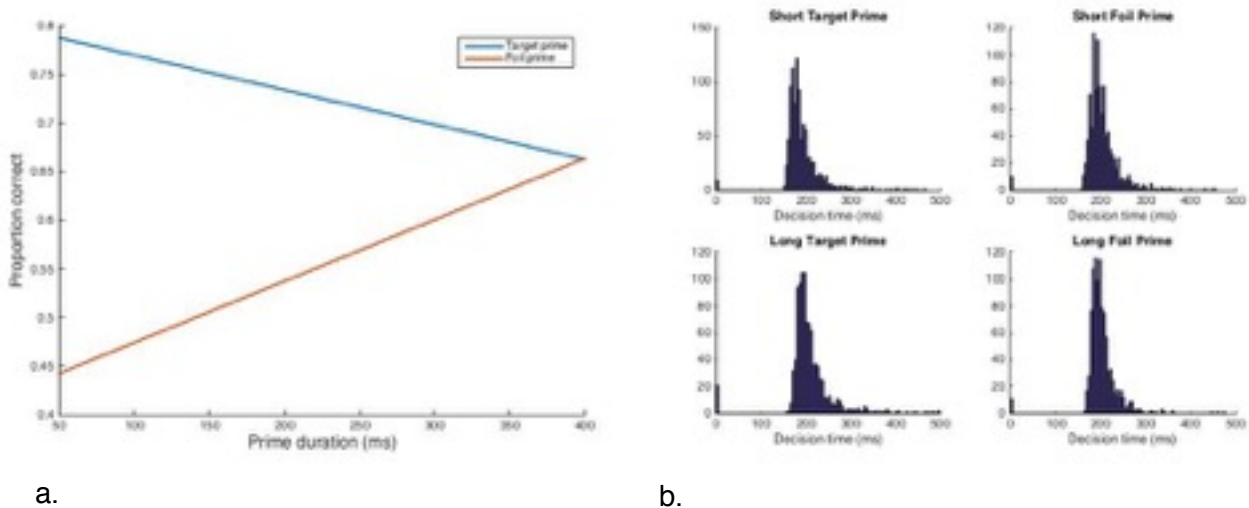

a.

b.

Figure 5. Stochastic visual encoding. Visual inputs are normally distributed with $\mu = 1$ and $\sigma = .3$. Input is drawn at each stage of visual presentation. Proportion correct vs. prime duration is shown in (a). 100-bin histograms of decision time (as defined by the shortest latency) for 4 conditions are shown in (b). Short primes are 50 ms, and long primes are 400 ms.

The next simulation exploits item differences in visual encoding. This rests on the idea that there may be systematic differences in the ability to visually encode different words. In these simulations, visual inputs for target and foil are drawn from a normal distribution with $\mu = 1$ and $\sigma = .3$. These visual inputs are then used for each presentation of target and foil, respectively. Results are shown in Figure 6.
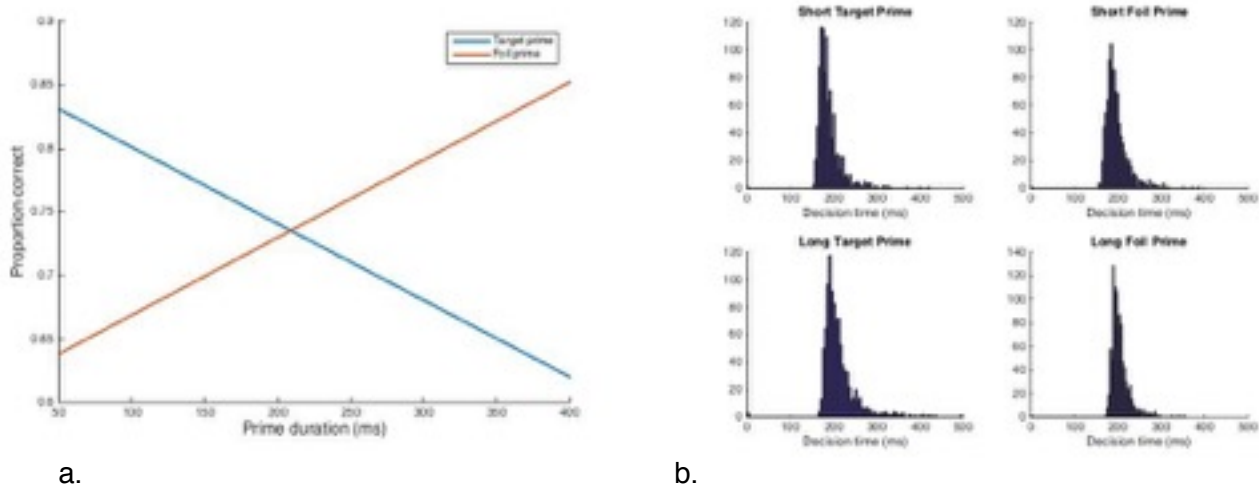

a.
b.

Figure 6. Item differences as realized through encoding variability. Visual inputs are normally distributed with $\mu = 1$ and $\sigma = .3$. Target and foil inputs are locked for each respective presentation throughout a given simulation trial. Proportion correct vs. prime duration is shown in (a). 100-bin histograms of decision time (as defined by the shortest latency) for 4 conditions are shown in (b). Short primes are 50 ms, and long primes are 400 ms.

4. Discussion

In this text, I have explored several sources of perceptual variability. I've considered the value of modeling each source of variability, considering 1) whether it produces non-extreme accuracies so as to overcome the need for a stochastic race-detection process and 2) whether it aids in modeling reaction times in repetition priming experiments. First, I examined item differences, the idea that there may be systematic differences in connection weights of different words, based on the extent of one's experience with each word. The results, shown in Figure 3,

suggest that modeling orthographic to lexical-semantic item differences is not a valuable step to take; the effects are simply not strong enough to warrant the added model complexity. Next, I found that visual to orthographic item differences could produce non-extreme accuracy in the characteristic crossover pattern in an accuracy vs. prime duration plot, shown in Figure 4a. However, these item differences do not produce the desired right-skew in decision times, as is shown in Figure 4b. These results suggest that, in certain circumstances, in may be valuable to consider item differences as realized through visual to orthographic connection weights, but that these item differences are not sufficient to achieve our goals of modeling reaction times. Next, I considered two forms of visual encoding variability; the first considers variability at the level of a stimulus flash, the second considers variability at the level of individual words. Visual encoding variability produced strongly variable target and foil latencies, so as to produce significant variation in performance for a given prime duration. When stochasticity was limited to one visual-display block at a time, it became clear that majority of the effects seen were held in the choice display block. Adding stochasticity to the visual choice block alone was sufficient to produce non-extreme, behaviorally-relevant accuracy data without any attempt to best fit the parameters. Additionally, this method produced a small right-skew in decision times. These results suggest that the choice display block is a very sensitive part of the model; differences in visual input between foil and target choices can significantly bias the outcome of the horse race. This implies that modeling variability in the choice block is a valuable course of action. I have only scratched the surface here; it is clear that variations in target/foil visual input in the display block could be modeled more explicitly in terms of selective attention, right/left biases, etc. These methods may be preferable to the current method of assuming a stochastic race-winner detection process, as they also produce variable decision times.

5. Comments and other work

Outside of my work, correspondence was found between the latencies produced by nROUSE and the drift rates of a model using diffusion to describe evidence accumulation toward target/foil choices. This correspondence was linked across 4 experimental conditions (foil/target prime, and short/long duration), in both same/different and forced choices tasks, for a total of 8 conditions. This is evidence that the nROUSE model naturally produces meaningful latency values.

Additionally, fits to individual data were made. Parameter values were for the most part adopted from prior fits to population data. However, the inhibition parameter I and the noise parameter N were allowed to vary for individuals, and an individual attention parameter was introduced. A sample of fits from this work is shown below, in Figure 7. As is clear, some fits are better than other. Formal statistical analysis of these fits remains to be done, and further work may allow for better individual-subject fitting than has currently been done. However, these individual fits are certainly a step in the right direction for the nROUSE model.

It seems likely that success in fitting nROUSE reaction times will come at the single-subject level. The success is fitting individual accuracy data leaves open the possibility of fitting a component of individual reaction times with the latencies produced by the nROUSE model. The success in producing variable latencies with different forms of stochastic processes suggests that individual are feasible.

Further plans exist at the level of behavioral experiments. An experiment has been designed to measure the visual input noise at different stages of processing, using visual gabor patches as stimuli. Behavioral measurements of internal and external noises provide a constrained way to estimate subject- and individual-level noise parameters within the model, which can then be used to fit reaction time data at the subject- and individual-levels.
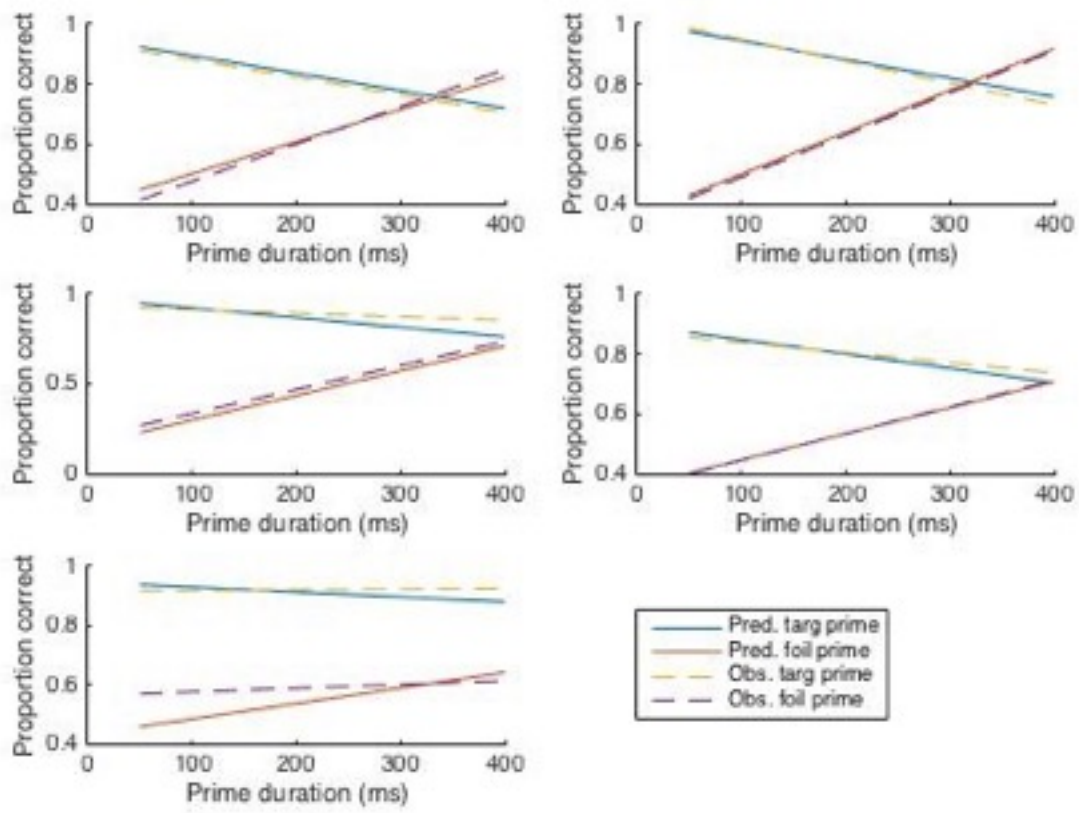
Figure 7. Sample of single-subject fits. 3 parameters are fit to single-subject data: attention, inhibition, and noise.