# Using a Neural Network Model with Synaptic Depression to Assess the Dynamics of Feature-Based Versus Configural Processing in Face Identification

**Cory A. Rieth (crieth@psyc.umd.edu)**
Department of Psychology, 1147 Biology/Psychology Building
College Park, MD 20742 USA

**David E. Huber (dhuber@psyc.umd.edu)**
Department of Psychology, 1147 Biology/Psychology Building
College Park, MD 20742 USA

**Abstract**

Accounting for the finding that brief prime durations facilitate perception of immediate word repetitions whereas long prime durations are detrimental, Huber and O'Reilly (2003) proposed a neural network model in which the unwanted effects of perceptual persistence are counteracted through activity dependent synaptic depression. Rieth and Huber (in prep) found similar results with immediate face repetitions, manipulating featural versus configural processing by means of face inversion. We extend the neural network model to face perception and account for individual differences by assuming some participants perform the task on the basis of feature identification, corresponding to the second layer of the 3-layer network, whereas other participants perform the task on the basis of configural identification, corresponding to the top layer. Under these assumptions, the model is used to describe the dynamics for each type of processing, with the resultant parameters revealing that configural identification integrates information at a faster rate than feature identification.

**Keywords:** Face Perception; Neural Network Modeling; Immediate Priming

## Introduction

One of the most highly specialized human perceptual abilities is the perception of faces. Without any conscious effort we are able to perceive, evaluate, recognize, and remember a large number of faces in any given day. This specialized type of expertise is often assumed to result from the rapid perception of configural aspects of faces, such as the distance between the eyes, nose, and mouth (Leder & Bruce 2000), although there is some debate as to the specifics of configural representation (Rakover, 2002).

The reported simulation studies do not specify which features are used to define face configuration or how face configuration is specifically calculated. However, assuming that feature detectors feed into configural detectors in a perceptual cascade, we describe the dynamics for each type of processing. We report evidence that configural face integrate information more quickly than feature detectors, with configural identification occurring based on partial feature information.

Huber, Shiffrin, Lyle, and Ruys (2001) developed an immediate priming paradigm that is particularly useful for assessing the dynamics of perceptual activation. Using visually presented words, they observed that briefly presented primes facilitated identification of an identical target whereas primes that were actively processed for several seconds were actually perceived less well. Weidemann, Huber, and Shiffrin (2005) modified the paradigm, demonstrating that these effects can be found simply as a function of prime duration. Rieth and Huber (in prep) applied this paradigm to the case of identification of novel faces, and these results serve as the focus of the reported simulation studies.

Huber and O'Reilly (2003) developed a dynamic neural network to account for the reversals in the direction of priming as a function of the extent of prime processing. Most human cortical cells produce a transient behavior termed synaptic depression, in which critical resources of the synapse are lost as a function of recent activity, resulting in greatly diminished signaling to receiver cells (Tsodyks & Markham, 1997). It is important to note that other biological mechanisms (e.g., calcium currents) also produce activity driven suppression, resulting in suppression of the entire cell rather than specific synapses. The results under consideration only included stimuli that were dissimilar or identical and, therefore, the reported simulations cannot speak to the difference between synapse specific depression versus depression of the entire cell. In any case, Huber and O'Reilly (2003) included such dynamics in a 3-layer neural network, proposing that activity driven suppression (e.g., synaptic depression) serves to dampen the unwanted effects of perceptual persistence, thereby clearing the way for subsequent perceptual items, at the possible expense of highly similar or identical items.

For the case of word identification, they assumed that layer 2 of the network processes orthographic information whereas layer 3 processes lexical-semantic information (layer 1 represents visual input, coding for different spatial positions). The model successfully accounted for patterns of results across the different priming manipulations and masking conditions by allowing each layer of the network to process information at its own rate, resulting in different degrees of persistence and inhibition within each layer. In accounting for word priming data, higher layers of the network were found to integrate more slowly.

One way to more precisely address the issue of dynamics at different layers of this perceptual cascade is by means of experimental manipulations that selectively harm one layer of processing. Face inversion is a particularly good example in that the exact same stimuli can be used for upright and inverted faces. Specifically, there is evidence that inverted faces are identified based on feature information whereas upright faces are identified based on configural aspects (Leder, & Bruce, 2000). Furthermore, studies with Event-Related Potentials (ERPs) find that inverted faces have an increased and delayed N170 component (Rossion et al., 1999). In accounting for immediate word repetition ERP results, Huber, Curran, O'Reilly, and Woroch (submitted) assumed that layer 3 of the network is the neural generator responsible for N170 effects. Therefore, in extending the model of Huber and O'Reilly (2003) to the case of face perception, it follows that layer 3 corresponds to configural processing, and that changes in the behavioral dynamics with face inversion are due to changes to this layer of processing.

In producing an account of face identification, we adopt the same 3-layer perceptual cascade model not because we believe the same brain areas are involved in words and faces, but rather because the functional form of different domains of visual expertise should be similar, with both face and word identification forming progressively complex representations and activity driven suppression in order to minimize temporal source confusion. By comparing resultant parameters between the face and word networks we can begin to specify structural and representational differences.

## Immediate Face Repetition

### Method

The modeled experiment (Rieth & Huber, in prep) was similar to that reported by Weidemann, Huber, & Shiffrin (2005) except that faces were used rather than words, and the orientation of the primes was modified from two vertical prime presentations to two horizontal presentations, (see Figure 1). In order to normalize the size and layout of the faces and to obtain a large number of unique faces to ensure that faces were not repeated over the course of the experiment, the FACES computer program was used to create 1,000 different faces. Half the trials presented upright prime faces and upright target faces, whereas the other half of trials presented inverted prime faces and inverted target faces. In the sequence of events, two identical prime faces were immediately followed by a briefly flashed target face. The prime face was either identical to the target face or identical to the incorrect foil face. Primes were presented for durations of 17, 50, 150, 400, or 2000 milliseconds. Target durations were determined separately for each participant by testing four different target durations intermixed across a series of trials, and then selecting the target duration that yielded performance closest to 75%. This was done to place target perception at threshold and accuracy in its most sensitive range. The tested target durations were 33, 50, 67, or 83 ms, although a target duration of 100 ms was possible if still under 75%. Target face presentations were immediately followed by a pattern mask. The duration between onset of the target face and offset of the pattern mask was fixed at 500 ms. Participants were given trial by trial feedback in order to discourage strategies in relation to the prime faces, and they were furthermore instructed that there could be no effective strategy because the prime face was just as likely to be identical to the wrong choice as the correct choice.

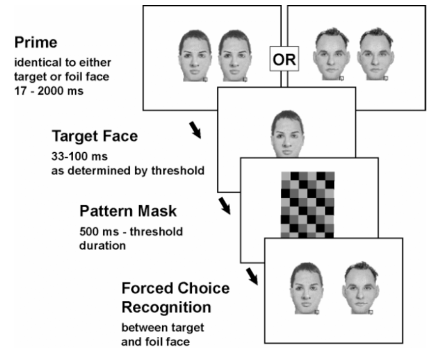


Figure 1: Sequence of events for the face priming experiment. All faces were male in appearance and included various forms of facial hair and hairstyles.

### Results

A first experiment ran 28 participants using only upright faces and a second experiment (see Figure 2) ran 40 participants with both upright and inverted faces, with the results for upright faces replicating the first experiment. The qualitative pattern of correct responses as a function of prime duration and prime type (foil primed or target primed) was similar to word priming except that the transition from priming benefits to deficits occurred at a somewhat slower pace as a function of prime duration. For words, the largest deficit for the foil prime condition occurs at approximately 50 ms, however in faces the largest deficit did not occur until 150 ms. For both words (not shown) and faces, it is important to note that the crossover between the target and foil primed conditions is related to, but not identical to the point at which the primes are identified.

Individual differences were investigated with a split half analysis; selecting participants who had target duration thresholds of 33 or 50 ms (low threshold group) versus those that required 83 or 100 ms (high threshold group). For participants who had a low threshold, the pattern was more similar to the word priming results, with a full crossover in the difference between the target and foil primed conditions as a function of prime duration. For the high threshold participants, the suppression of prime faces after extended prime durations was apparently not as pronounced, only managing to eliminate, but not reverse the priming effects.

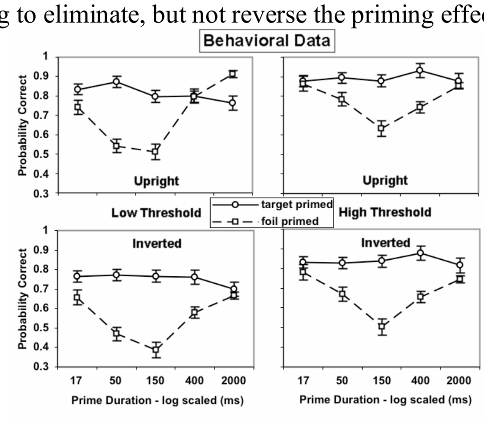


Figure 2: Behavioral Results. Error bars are +/- one SEM.

Next we consider the effect of face inversion. For the low threshold group there was a sizable effect of inversion such that performance was worse, and, furthermore, the crossover as a function of prime duration failed to fully emerge. Surprisingly, there was no apparent effect of face inversion for the high threshold group, with performance roughly equivalent and a lack of crossover for both upright and inverted faces. In summary, only for the low threshold group with upright faces was there a crossover pattern.

Considering that the low threshold group required shorter target durations for threshold performance, and considering the lack of inversion effects for the high threshold group, this suggests that low threshold participants processed upright faces in a configural manner, whereas the high threshold participants processed faces on the basis of individual features, even when those faces appeared in their proper upright orientation. In general, identifying faces on the basis of individual features (e.g., hair styles, nose size, etc.), is not viewed as particularly effective, although such a strategy may have been tempting in this experiment considering the heterogeneity in the features used to comprise these particular faces. Of course these claims are generalities, and it is unlikely that such strategic differences should exactly align with a split half analysis of participants. Nevertheless, assuming that in general the low threshold participants processed faces in a configural manner when viewing upright faces, the results of the split half analysis are particularly useful for specifying the dynamics of configural and featural face processing.

As described next, we applied the neural network model Huber and O'Reilly (2003) to these data, providing further support for the claim that the high threshold participants engaged in feature based identification, regardless of face orientation, whereas the low threshold participants engaged in configural processing for upright faces, but necessarily fell back upon feature processing for inverted faces. Furthermore, for the case of the low threshold group in which there were inversion effects, the resultant best-fit parameters yield a quantitative measure for the speed of processing for configural as compared to feature representations.

## A Neural Network Model of Face Processing

The reported simulation studies used an artificial neural network consisting of the three layers seen in Figure 3. The layers in the model as applied to face processing, consist of a lower visual layer that encodes sensory input from different regions of the visual field, followed by a middle layer that encodes face features, and finally a top layer that encodes specific face configurations. We did not employ learning in the construction of the model. The goal of the model is to capture the dynamic properties of face perception, rather than specify the particular representations.

Because similarity was not manipulated in the experiment, a localist representation was employed with full (parameterized) connection strengths between the units encoding for a particular face and connection strengths of 0.0 between units encoding for different faces. Each unique face was assigned a different representative featural unit and a different representative configural unit. Because the visual layer encodes for different areas of the visual field, each face was assigned a different visual unit for each location presented within the sequence of events. These units mapped to the same featural unit regardless of location. This many to one mapping produces temporal integration across different presentations of the same face.

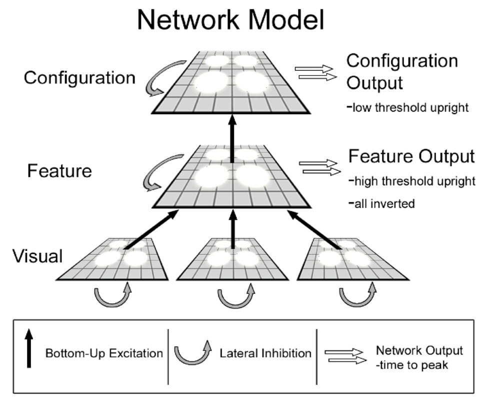


Figure 3: Neural network organization.

All-to-all lateral inhibition was utilized within each layer of the network to dampen excessive activation. This inhibition produces masking effects for items presented in the same visual location. Application of the model to these novel faces did not include feedback between layers under the assumption that novel faces look somewhat similar to known faces, but the overall effect of feedback is negligible (for familiar words, feedback between the lexical-semantic layer and orthography was set to .25, and there was no feedback to the visual layer). In other words, a novel face may look like the configuration of a known face, but that known face is just as likely to provide top-down support for the incorrect features as the correct features. In any case, future research will manipulate item familiarity and assess how these manipulations relate to connection strengths within the network.

### Activation Dynamics

The model builds upon the LEABRA framework of O'Reilly and Munakata (2000). Individual simulated neurons represented activity through a rate code (i.e., probability of spiking), and, as such, can be viewed as a stand-in for entire assemblies of neurons that have similar inputs and outputs. Further simplifying from the known biology, these simulated neurons are "point neurons", and do not explicitly include temporal delays or non-linearities that may result from transfer of charge along dendrites and axons as in compartmental artificial neurons. Membrane potential in these simplified neurons is updated as a function of the excitatory inputs to the neuron, lateral inhibitory connections, and leak currents as in Equation 1.

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

The change in membrane potential per unit time ($\Delta v_i^n$) for neuron $i$ in layer $n$ is updated as a function of the previous time step's voltage ($v_i^n$), the weighted output of the neurons from the layer below ($w_{ij}o_i^{n-1}$), the leak parameter ($L$), lateral inhibition ($I$), the output of other neurons within the layer ($o_k$), and, most importantly for the current situation, the value $S_n$ representing a proportion of change for each millisecond, which is used to capture different rates of integration for different layers. Activity in these neurons is thresholded, with no response occurring when membrane potential fails to achieve the requisite value ($\theta$). This thresholded membrane potential is additionally scaled by a dynamically varying term (see Equation 2) that captures the resources available at the synapse ($a$), with the product of these two terms providing the output ($o$) for any particular neuron.

$$o = \begin{cases} (v - \theta)a & v > \theta \\ 0 & v \leq \theta \end{cases} \tag{2}$$

As a function of recent output ($o$), synaptic resources are depleted according to Equation 3, thereby simulating suppression from synaptic depression. As with membrane potential, the same rate of integration parameter $S_n$ is used for updating synaptic resources for the neurons of layer $n$. This parameter represents a general metabolism for the activity within a layer, and is therefore applied to both membrane potential update as well as synaptic resources update. In this manner the rate of synaptic depression is kept proportional to the rate of temporal integration (although note that synaptic depression is itself a very dynamic property, with above threshold membrane potential the driving force behind synaptic depression). The parameter $R$ is the rate of recovery for synaptic resources, and $D$ is the rate of depletion for synaptic resources, with each of these fixed to the same values for all neurons. By multiplying thresholded membrane potential by a dynamically varying value for synaptic resources, as well as the influence of connection strengths, which are set to 0 or full strength (1.0 for input to the first layer, and free parameters for connections between the other layers), the magnitude of excitatory input to receiver neurons reflects: pre-synaptic activity, available synaptic resources and connection strength.

$$\frac{\Delta a_i^n(t)}{S_n} = R(1 - a_i^n) - D o_i^n \tag{3}$$

The model was run in time steps of 1 millisecond, and input to the visual layer was set at 1.0 when stimuli were presented and 0.0 otherwise. As the resultant activation travels up the perceptual cascade, the effect of temporal integration (Equation 1) produces persistence, with the duration of persistence increasing with each additional layer. Furthermore, with the output of each layer subject to synaptic depression, the response of each layer above the visual layer reaches a peak level and then falls off, with higher layers peaking at progressively later delays.

Because participants are placed at their perceptual threshold (i.e., target duration set such that performance is only 75% with forced choice testing), we assumed that explicit identification of the briefly flashed target face was not possible, and, therefore, performance was based upon partial information in a matching process to the choice alternatives. Such partial information could be assessed through the level of remaining activity at the time of the choice alternatives, although this measure has the undesirable characteristic that very long target durations actually result in worse performance due increases in synaptic depression. Instead, Huber and O'Reilly (2003) assumed a fluency measure of partial information in which the choice alternative that reaches its peak response first is chosen. This decision rule is essentially to choose the face that "leaps out" at the participant first. Furthermore, this decision rule is roughly equivalent to a horse race model of forced choice testing (although the racers are not fully independent due to lateral inhibition), and, as Huber and Cousineau (2004) demonstrated, such a horse race model is capable of capturing both correct and error reaction time distributions for these immediate priming tasks with forced-choice testing.

In the current application, the model was run in a deterministic manner with the difference between the time-to-peak response of the target versus the foil converted into an accuracy measure through a logistic sigmoid function with the parameter N (see Equation 4, $T_F$ and $T_T$ are the time to peak of the foil and target neurons), with N inversely related to amount of noise (higher values of N result in better performance). This rescaling of the difference between the time to activate each choice alternatives is envisioned as resulting from processing noise in the update equations, although exact specification awaits future modeling studies.

$$p(c) = \frac{e^{N(T_F - T_T)}}{1 + e^{N(T_F - T_T)}} \tag{4}$$

With this decision rule, briefly presented prime faces result in persistent activity, providing a head start for the primed choice face. This explains why performance is enhanced in the target primed condition and harmed in the foil primed condition for briefly presented primes. However, following excessively processed primes (i.e., long prime durations), the extent of this persistent activity is greatly diminished due to synaptic depression. Beyond this reduction in the amount of persistence (which would only reduce, but not reverse priming effects), synaptic depression also serves to slow down the re-activation of a face that has recently been viewed. This sluggishness to respond is not due to any single layer in isolation, but is instead emergent from the dynamics of processing across layers. Thereby, this "disfluency" produces a full reversal in the direction of priming.

In order to fully understand the pattern of model behavior, the role of inhibition is important. Related primes have a direct effect on the decision process through persistence and synaptic depression, but in addition there is an indirect effect due to inhibition, which affects target performance regardless of priming condition, primarily by diminishing how much additional activation is accrued in response to the briefly flashed target. Primes reach their maximal activation around 150 ms for the reported parameters, and so the detrimental effects of inhibition are maximized for prime durations of 150 ms. This explains why overall performance initially decreases with increasing prime duration, up to approximately 150 ms, but then overall performance increases for even longer prime durations.

## Mapping Layer Responses to Strategies

This model is only intended to capture the dynamics of perception and does not attempt to explain attentional and response factors that may vary as a function of the task. We simplify the situation by assuming performance based on feature identification is achieved through the fluency response (i.e., time-to-peak) for the output of layer 2, whereas performance based on configural identification is achieved through the fluency response for the output of layer 3. Additional work is required to fully integrate this model into a general cognitive architecture, thereby explicating how the decision process can be reformulated to selectively attend to one type of fluency response versus the other. With this simplification, the model was applied to the face priming data by assuming low threshold participants used layer 3 fluency for upright faces but layer 2 for inverted faces. Furthermore, the high threshold group was assumed to use layer 2 fluency for both upright and inverted faces. Under these assumptions, application of the model to these data is used to ascertain the relative speed of processing ($S_n$) for each layer.

We modeled the face priming results with the inhibition, leak, depression, recovery, and threshold parameters that Huber and O'Reilly (2003) used for this same paradigm with immediate word repetitions. Even though brain areas involved in face processing may be different in important ways from the areas involved in the perception of words, these parameters represent general properties of neurons and should not vary significantly throughout the cortex, particularly in two areas both involved in visual perception. The values of these fixed parameters were: inhibition, $I = 0.3$, threshold, □ = 0.15, leak, $L = 0.15$, first layer rate of integration, $s_1 = .054$, depression, $D = 0.324$, and recovery, $R = .022$. To capture the face perception data, connection strengths from the first to second and second to third layer, the integration speed of the second and the third layer, and the noise term in the logistic function were freely varied.

## Simulation Results

As seen in Figure 4, the model successfully captured the qualitative patterns of results as a function of prime duration, upright versus inverted presentation, and also for the two groups of participants. A chi square measure of goodness of fit with one degree of freedom was computed for each of the 40 conditions as described in Batchelder & Riefer (1990), with 320 data points per condition. The median chi square was 2.07. The model was not statistically different from the behavioral data in 24 conditions (□ = .05). The best fitting parameters for the low threshold subjects were: $C_{12} = 1.20$ (connection strength from the visual layer to the feature layer), $C_{23} = 1.09$ (connection strength from the feature layer to the configural layer), $N_{UP} = .022$ (noise term for upright faces with the output of layer 3), $N_{INV} = .043$ (noise term for inverted faces with the output of layer 2), $s_2 = .014$ (speed of integration for the feature layer), and $s_3 = .021$ (speed of integration for the configural layer). For the high threshold subjects the best fit parameters were: $C_{12} = .754$, $N_{UP} = .062$ (note these noise terms are used with layer 2 output), $N_{INV} = .040$, and $s_2 = .014$ (there was no need to include layer 3 parameters for the high threshold group).

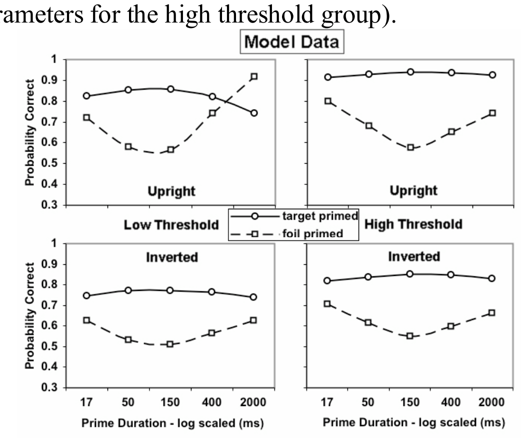

Figure 4: Neural network results.

## Discussion

Comparing Figures 4 and 2, this modeling investigation provides a reasonable approximation of the behavioral results under the assumption that different participants adopted different identification strategies corresponding to feature identification versus configural identification. Although the quantitative fit is not exact, the model captures all the qualitative aspects in the behavioral data.

Besides an existence proof that the model assumptions can adequately explain the observed data, another use of computational modeling is interpretation of observed data in terms of the underlying psychological or biological parameters. In the current situation, one of the goals of this modeling exercise was to compare the rate of processing ($S_n$) for the feature layer as compared to the configural layer. The result that a crossover between the target and foil primed conditions as a function of prime duration only occurred for the low threshold group with upright faces is of particular importance under the assumption that configural processing is only used for upright faces, and the model captured this effect by setting the rate of processing for the configural layer to a higher value (.021) than for the feature layer (.014). For Huber and O'Reilly's (2003) simulations with words, the speed of integration for lexical-semantic processing (layer 3) was much slower than orthographic processing (layer 2), whereas the opposite was true for the current application to faces. Assuming the model provides a reasonable approximation of the time course of perceptual processing, this is an important. This suggests that words are processed in a componential manner, with more abstract representations awaiting full verification from more concrete representations, whereas faces are processed more holistically, with configural representations essentially "jumping the gun" in their identification based on only preliminary feature information.

There may be multiple causes for this difference between componential and holistic processing when comparing words and faces. Lexical identification depends not only on the relative positions of letters within a word, but additionally on the exact identification of those letters. In contrast while feature identification undoubtedly plays some role in face perception, it is not clear that configural face identification requires complete identification of the features (e.g., Bob's nose or Sally's eyes), and may only rely upon identification of feature classes (e.g., a nose versus an eye). When considering the different degree of componentiality comparing words and faces, it may be advantageous to wire face processing with more rapid configural processing, whereas this would be detrimental for word processing.

Beyond the degree of componentiality of words and faces, faces have been a dominant and important visual stimulus for the history of our species, and it is clear that we utilize special mechanisms for their perception. In contrast, writing has existed for less than 10,000 years, and is perhaps more reliant upon ontological adaptation. This specialization for rapid identification of configuration may therefore result from evolutionary mechanisms, providing a kind of expertise that cannot be equaled merely through a lifetime of learning.

Providing additional support for our claims, split half analyses based on threshold target duration with words failed to produce the differences that we see with faces. According to the model, lexical-semantic processing occurs at a slower rate than orthographic processing and so layer 3 does not infer much additional suppression for words. As such, a strategy to focus on the lexical-semantic or orthographic representation does not produce much of a difference in terms of the crossover between target primed and foil primed. In contrast, the configural layer effectively "runs faster" for faces and therefore provides a healthy dose of additional suppression. As a result, face priming effects depend more heavily on the performance strategy, with more rapid identification and supression when attending to configuration.

An alternative argument could be made that the division of participants into low and high threshold groups is not based on response strategy, but on individual differences in face perception ability. The key difference may be that while the target duration is tailored for each individual's rate of face processing, the various prime durations are not specifically tailored for each individual. This would result in apparently more rapid priming effects (i.e., as a function of prime duration) for the low threshold group. Such an explanation is appealing, but fails to explain the interactions with face inversion, and also fails to explain why we do not see similar individual differences with word priming where there is even a larger range of individual differences in the required target flash durations to achieve threshold performance.

In future work we will seek converging evidence for this interaction between prime duration and configural versus featural processing by using other techniques for manipulating configurality, such misalignment (Young, Hellawell, & Hay 1987). In addition we will see how these priming effects change as participants become progressively more familiar with a limited number of previously novel faces. It is hoped that with further experimentation and further refinement of the neural model, we can develop a more complete and accurate understanding of face perception.

## Acknowledgments

Funding for this research was provided by NIMH grant number MH063993-04. We also thank the anonymous reviewers for their comments and suggestions.

## References

Batchelder, W. H. & Riefer, D. M. (1990). Multinomial processing models of source monitoring. *Psychological Review*, *97*, 548-564.

Huber, D. E., & Cousineau, D. (2004). A race model of perceptual forced choice reaction time. *Proceedings of the 25th Annual Conference of the Cognitive Science Society.*

Huber, D. E., Curran, T., O'Reilly, C, & Woroch, B. (submitted). Interpreting ERPs with a biologically-based neural network: The dynamics of integration and separation. *Journal of Experimental Psychology: Human Perception and Performance.*

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(3), 403-430.

Huber, D. E., Shiffrin, R. M., Lyle, K. B., & Ruys, K. I. (2001). Perception and preference in short-term word priming. *Psychological Review*, 108(1), 149-18.

Leder, H., & Bruce, V. (2000). When inverted faces are recognized: The role of configural information in face recognition. *The Quarterly Journal of Experimental Psychology*, 53A, 513-536.

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

Rakover, S. (2002). Featural vs. configurational information in faces: A conceptual and empirical analysis. *British Journal of Psychology*, 93, 1-30.

Rieth, C. A., & Huber, D. E. (in prep). The dynamics of face identification: Interactions between immediate face repetition and configural processing.

Rossion, B. Delvenne, J., Debatisse, D., Goffaux, V., Bruyer, R., Crommelink, M., Guerit, J. (1999). Spactio-temporal localization of the face inversion effect: an event-related potentials study. *Biological Psychology*, 50, 173-189.

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

Weidemann, C. T., Huber, D. E., Shiffrin, R. M. (2005). Confusion and compensation in visual perception: Effects of spatiotemporal proximity and selective attention. *Journal of Experimental Psychology: Human Perception and Performance,* 31, 40-61.

Young, A.W., Hellawell, D., & Hay, D.C. 1987. Configurational information in face perception. *Perception, 16,* 747-759.