

Temporally distinct neural coding of perceptual similarity and prototype bias

David Alexander Kahn

Department of Neurology, University of Pennsylvania,
Philadelphia, PA, USA



Alison M. Harris

Division of Humanities & Social Sciences, California Institute
of Technology, Pasadena, CA, USA



David A. Wolk

Department of Neurology, University of Pennsylvania,
Philadelphia, PA, USA



Geoffrey Karl Aguirre

Department of Neurology, University of Pennsylvania,
Philadelphia, PA, USA



Psychological models suggest that perceptual similarity can be divided into geometric effects, such as metric distance in stimulus space, and non-geometric effects, such as stimulus-specific biases. We investigated the neural and temporal separability of these effects in a carry-over, event-related potential (ERP) study of facial similarity. By testing this dual effects model against a temporal framework of visual evoked components, we demonstrate that the behavioral distinction between geometric and non-geometric similarity effects is consistent with dissociable neural responses across the time course of face perception. We find an ERP component between the “face-selective” N170 and N250 responses (the “P200”) that is modulated by transitions of face appearance, consistent with neural adaptation to the geometric similarity of face transitions. In contrast, the N170 and N250 reflect non-geometric stimulus bias, with different degrees of neural adaptation dependent upon the direction of transition within the stimulus space. These results suggest that the neural coding of perceptual similarity, in terms of both geometric and non-geometric representations, occurs rapidly and from relatively early in the perceptual processing stream.

Keywords: event-related potentials, perceptual similarity, neural adaptation, prototype effect, N170

Citation: Kahn, D. A., Harris, A. M., Wolk, D. A., & Aguirre, G. K. (2010). Temporally distinct neural coding of perceptual similarity and prototype bias. *Journal of Vision*, 10(10):12, 1–12, <http://www.journalofvision.org/content/10/10/12>, doi:10.1167/10.10.12.

Introduction

From searching for one’s car in a parking lot to finding a friend in a crowd, we are confronted daily with varying exemplars from a given visual category. How does the visual system represent this variety? Several perceptual models are built around the notion of a “stimulus space,” a representation of comparative similarity based on observers’ judgments or their classification of stimuli into groups. Within-class stimulus variation may be mapped along the dimensions of this space. Rectangles, for instance, can be described in terms of aspect ratio and area, and color defined by variation in hue, saturation, and brightness.

A number of psychological models have related stimulus spaces to behavioral measures of perceptual similarity. So-called “geometric” models postulate a direct correspondence between the two, defining similarity in terms of the metric distance between two stimuli within a representational space (Shepard, 1964; Torgerson, 1965). While such geometric models are successful in explaining a wide

range of behavior, certain perceptual properties of similarity violate these models (Holman, 1979; Krumhansel, 1978; Tversky, 1977). Notable is the violation of symmetry: while the ordering of a pair of stimuli should not alter their perceptual similarity in geometric models, this violation is frequently seen in practice. A classic perceptual example is that an ellipse is judged to be more similar to a circle than a circle is to an ellipse (Tversky, 1977). Often, such asymmetries suggest the existence of representational “prototypes,” which can be interpreted as stimulus-specific biases producing non-geometric distortions of otherwise geometric similarity spaces. Prototypes may be the result of long-standing perceptual experience or the local effect of context induced by stimulus frequency (Polk, Behensky, Gonzalez, & Smith, 2002). Current models of similarity account for perceptual asymmetries through the inclusion of both geometric and non-geometric properties. The “additive similarity and bias” model of perceptual proximity (Holman, 1979; Nosofsky, 1991), for example, incorporates both geometric and non-geometric effects by defining the perceptual “proximity” of two stimuli as the sum of metric stimulus distance and

stimulus bias, a term representing the stimulus-specific effects behind such asymmetries.

Supporting this distinction, studies of the neural representation of stimulus similarity have identified both geometric and non-geometric neural codes. A single-unit study of object perception demonstrated a correspondence between neural responsiveness in monkey inferotemporal cortex and the geometric organization of an abstract shape space, as derived from both behavioral and pixel-wise evaluations of similarity (Op de Beeck, Wagemans, & Vogels, 2001). Analogous geometric effects of similarity have been demonstrated in regions associated with object perception in humans using functional magnetic resonance imaging (fMRI; Drucker & Aguirre, 2009). Non-geometric similarity codes, in contrast, have been proposed to explain differential responsiveness to “prototypical” faces as compared to “distinctive” faces in fMRI (Loffler, Yourganov, Wilkinson, & Wilson, 2005).

Yet a great deal about the neural representation of perceptual similarity remains poorly understood. One major question relates to the dissociation of geometric and non-geometric effects at the neural level. While each of the studies cited above demonstrates neural correlates of either geometric or non-geometric encoding, no existing study has examined both types of effects concurrently. A second question is the time course of perceptual similarity effects: when in the perceptual processing stream do geometric and non-geometric coding of stimulus similarity occur? This latter question, extending to the temporal domain, speaks to the former by providing a non-spatial means of distinguishing these components of perceptual similarity.

In the present study, we investigated these questions using event-related potentials (ERPs). We hypothesized that geometric and non-geometric features of similarity would be evaluated during the time course of visual perception and focused upon several of the early perceptual and “face-selective” components of the evoked visual response. In our analysis, we examined four components of the ERP waveform previously associated with various stages of perceptual and mnemonic processing for faces. These include the P100, a marker of early visual processing (e.g., Di Russo, Martínez, Sereno, Pitzalis, & Hillyard, 2001), the N170 (occurring approximately 170 ms after stimulus onset), which is associated with perceptual encoding of the face (Bentin, Allison, Puce, & Perez, 1996; Itier & Taylor, 2004; Liu, Higuchi, Marantz, & Kanwisher, 2000; Sams, Hietanen, Hari, Ilmoniemi, & Lounasmaa, 1997), the P200, the positive component following the N170, and the N250, thought to reflect consolidation of perceptual representations into memory (Tanaka, Curran, Porterfield, & Collins, 2006). We used these components as elements of a temporal framework on which a neural model of geometric and non-geometric similarity effects could be evaluated.

We examined the sensitivity of this temporal framework to perceptual similarity by presenting faces varying in

identity between two endpoint faces. Sensitivity to perceptual similarity was assessed via neural adaptation: a reduction in neural response following repeated stimulus presentation (Grill-Spector & Malach, 2001; Henson & Rugg, 2003). Previous work has demonstrated neural adaptation of “face-selective” responses in ERP (Itier & Taylor, 2002; Jacques & Rossion, 2006; Kovács et al., 2006; Schweinberger, Pickering, Jentsch, Burton, & Kaufmann, 2002) and the related methodology of magnetoencephalography, or MEG (Furl, van Rijsbergen, Treves, Friston, & Dolan, 2007; Harris & Nakayama, 2007, 2008). However, few of these studies have tested for parametric variation of adaptation effects, and the measurement of geometric and non-geometric similarity effects are often confounded. For example, while studies of prototype representation may observe differential response to centrally located stimuli (e.g., Loffler et al., 2005), these effects may result from the tendency of prototypical stimuli to be more similar to other stimuli and thus produce neural adaptation.

To disentangle these effects, we used a “carry-over design” (Aguirre, 2007) in which a continuous stream of stimuli is presented with first-order counterbalancing. The resulting data permit measurement of the direct effect of each stimulus upon the amplitude of neural response, as well as the modulatory effect of one stimulus upon the next (e.g., neural adaptation). Geometric neural similarity is revealed in this context as a symmetric, parametric adaptation of ERP response proportional to the change in perceptual similarity. Non-geometric neural similarity, suggestive of explicit neural representation of a prototype or central tendency of the stimulus space, was modeled as an asymmetric modulation of the ERP response dependent upon the direction of stimulus transition.

Materials and methods

Subjects

Six right-handed subjects (3 women, 3 men) between the ages of 22 and 39 (mean age 29.5) with normal or corrected-to-normal vision participated in the study. All subjects provided informed consent under the guidelines of the Institutional Review Board of the University of Pennsylvania and the Declaration of Helsinki.

Stimuli

Two neutral faces (subtending $9.4^\circ \times 10.9^\circ$ of visual angle) adapted from the NimStim stimulus set (Tottenham et al., 2009), varying in eye and mouth identity, were used to create a linear morph, yielding five stimuli varying in 25% increments. (Since the actual images used for experimentation are not publishable, all figures use sample

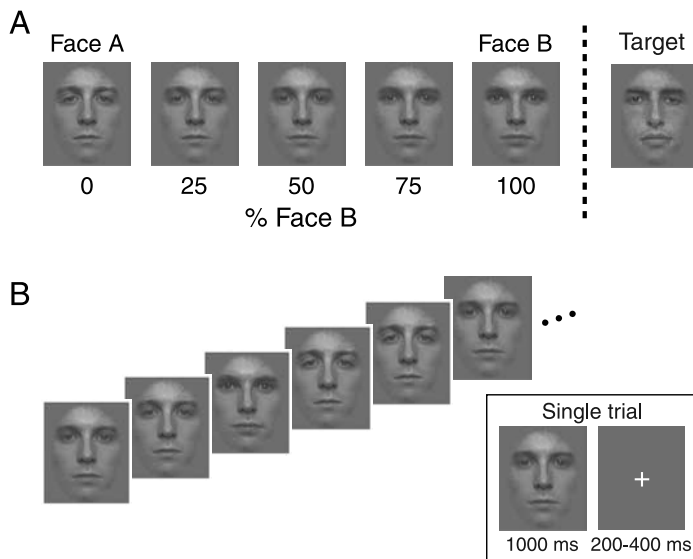


Figure 1. Sample stimuli and presentation. Representative sample stimuli are presented here as the actual stimuli used were not approved for publication. (A) The experimental stimuli consisted of five faces morphed in identity between two endpoint identities (Faces A and B) in 25% increments; subjects were not informed of the stimulus space arrangement. Subjects were instructed to monitor for the appearance of a target face (far right) whose identity was distinct from the morph axis. (B) Stimulus presentation. Stimuli were presented for 1000 ms with an ISI of 200, 300, or 400 ms, counterbalanced across trials using a type 1, index 1 sequence (Aguirre, 2007) with 18 elements.

morphs from a different stimulus set.) All faces (Figure 1A) were cropped of external facial features using the same selection boundary shape (ellipse, 3-pixel feathering) and set to grayscale bitmaps in Adobe Photoshop.

The similarity of the resulting face images was analyzed using a biologically motivated, multi-scale, Gabor-filter model of V1 cortex (Renninger & Malik, 2004). A multi-dimensional scaling (MDS) analysis of the computational similarity scores revealed that, as expected, the faces varied along a single dimension and had roughly equal spacing between the 5 stimuli (spacing between adjacent, nominal 25% morphs: 30%, 24%, 21%, 25%).

Behavioral assessment of stimulus similarity

A behavioral, reaction time study was used to confirm the monotonic ordering of the perceptual similarity of the stimuli along the face morph continuum. All subjects ($N = 6$) from the ERP study participated in the behavioral study several days following ERP data collection.

The 5 faces from the morph continuum were used as stimuli and presented side by side on a computer screen using the PsychToolbox (Brainard, 1997; Pelli, 1997) for MATLAB (Mathworks, Andover, MA). Subjects were

instructed to respond with a button press to indicate if the pair of faces were the same or different (buttons indicating same or different were randomized to right or left across subjects). Each trial consisted of a side-by-side face presentation lasting until the subjects responded with a button press, followed by a 250-ms inter-trial interval. Runs consisted of 640 trials, with breaks occurring every 40 trials. “Same” trials, in which the face identity was the same, occurred with equal frequency as “different” trials. Within the “different” trials, the metric distances ($\Delta 25$, $\Delta 50$, $\Delta 75$, $\Delta 100$) along the morph continuum occurred with equal frequency.

For each different face pair for each subject, the inverse of the median of correct reaction times was found and entered into a distance matrix for multi-dimensional scaling (MDS) analysis (Kruskal & Wish, 1978). MDS analysis for each subject was performed for each subject using the MATLAB `cmdscale()` function. Coordinates were centered about the 50% face for each subject, and then averaged across subject to yield estimates of stimulus placement. The first dimension of the MDS estimate was retained.

ERP stimulus presentation

Each run consisted of 648 trials; each subject underwent 3 consecutive runs for a total of 1944 stimulus presentations. Each trial consisted of a stimulus presentation for 1000 ms, followed by an ISI of 200, 300, or 400 ms (counterbalanced across trials). Stimulus order was determined by a first-order, counterbalanced, $n = 18$, type 1, index 1 sequence (Aguirre, 2007). An 18-element sequence was required to counterbalance the 6 stimuli (5 morphs and 1 target) crossed with the three durations of ISI that could follow each stimulus. During the ISI, a central white fixation cross was presented on the same mean gray background surrounding the stimuli. Subjects were instructed to respond with a button press to the occurrence of a target face from outside the morph continuum (Figure 1A, far right). Subjects were trained on a simplified version of the task immediately prior to the experiment to ensure accurate identification of the target face. Target trials and trials immediately following target presentations were excluded from the main analyses.

Stimuli were presented using EPrime 2 (Psychology Software Tools) on a Dell 24-inch LCD display situated 100 cm from the subject at eye level. Task responses were also collected through EPrime 2. To obtain “sensors of interest” for experimental analysis, after the main experiment subjects completed a short “localizer” experiment with faces, houses, and everyday objects (100 exemplars each) randomly interleaved. Stimuli in the localizer were presented on a white background with a black fixation cross ($9.2^\circ \times 7.7^\circ$ visual angle) for 300 ms (ITI jittered between 900 and 1100 ms); subjects were instructed to passively view the stimuli.

ERP data collection

Data collection was performed on a BioSemi ActiveTwo system (<http://www.biosemi.com/products.htm>) with 128 active electrodes with sintered Ag–AgCl tips in fitted headcaps. Evoked brain potentials were digitized continuously at a sampling rate of 512 Hz with default low-pass filtering at 1/5 of the sampling rate (http://www.biosemi.com/faq/adjust_samplerate.htm). Two additional electrodes with a 4-mm sintered Ag–AgCl pallet were also placed bilaterally on the mastoids as references for data import (<http://www.biosemi.com/faq/cms&drl.htm>). Electrical offsets were verified to be between -20 and $20 \mu\text{V}$ for every channel prior to data collection.

ERP pre-processing and analysis

Data were processed offline using the EEGLAB toolbox (Delorme & Makeig, 2004) for MATLAB. Sensors were selected for analysis using a “sensor of interest” (SOI) approach (Liu, Harris, & Kanwisher, 2002), via a point-to-point t -test comparing face and house conditions in the “localizer” scan. Significant channels for each subject were identified within the N170 and N250 latency ranges, and group channels (Figure 2A) used for subsequent analysis were selected if they were identified as significant in a majority of subjects (4 out of 6). Group average

waveforms across all non-target trials for each sensor can be found in [Supplementary Figure 1](#).

All data for each subject were saved from BioSemi ActiView and imported by run directly into EEGLAB. Mastoid channels were indicated as references to EEGLAB upon import and excluded; data were re-referenced immediately to the average signal of all 128 cranial channels. Data were epoched to a time window of 700 ms (100-ms pre-stimulus onset and 600-ms post) and baseline corrected (100-ms pre-stimulus onset). Trials containing artifacts (e.g., eye blinks) were identified and removed automatically using a $\pm 100 \mu\text{V}$ threshold (average rejection rate across subjects for trials used in the main analysis was 16.7%, with a range of 5.3%–38.8%).

ERP components of interest were identified for each subject individually using data averaged over all non-target conditions across the “sensors of interest” defined at the group level (Figure 2B). The previously described P100 and N170 were defined on the basis of latency and direction of deflection, while the N250 was defined by the comparison of target and non-target faces (Tanaka et al., 2006). Inspection of our results also revealed a meaningful deflection between N170 and N250, here called the P200. For each subject’s grand average waveform, the time points of the local minima (for N170 and N250) and local maxima (for P100 and P200) were identified within search windows (P100: 125–175 ms; N170: 175–225 ms; P200: 225–275 ms; N250: 300–350 ms) and used as centers of the respective components for that subject. For each

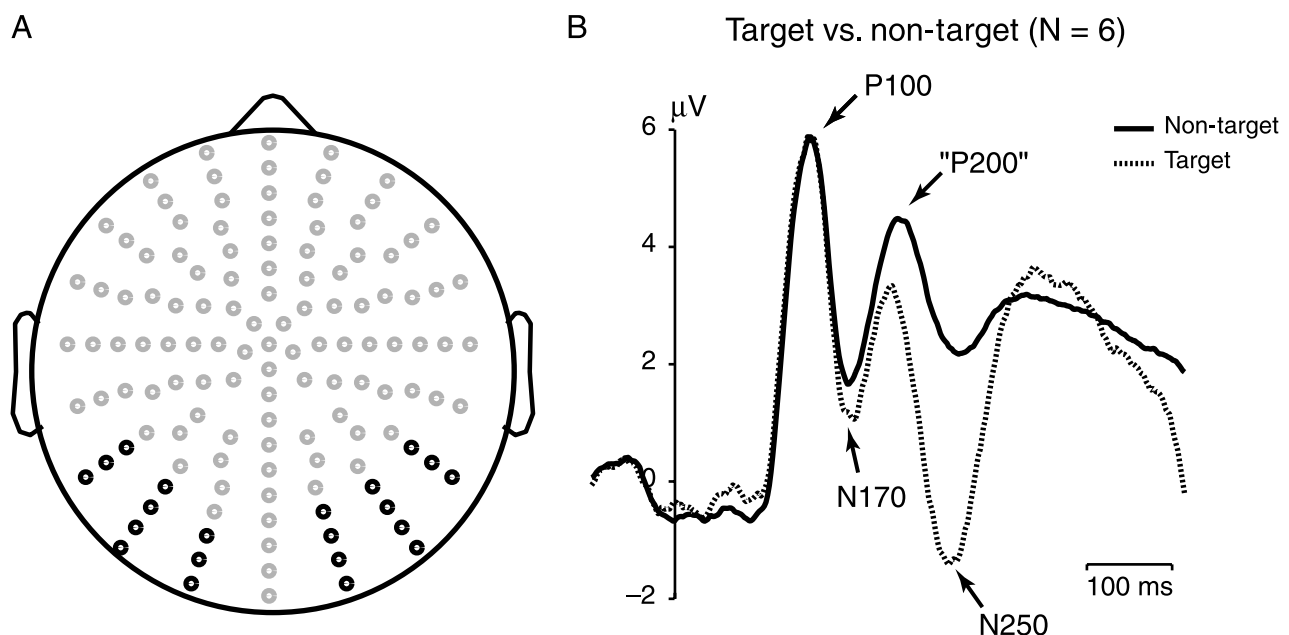


Figure 2. ERP sensor of interest (SOI) selection and component definition. (A) Twenty-one face-selective (black dots) SOIs were selected across subjects using an independent localizer task (Face > House). (B) Component identification. Grand average waveforms ($N = 6$) comparing the response to trials in which the target face was presented and all non-target trials. P100 and N170 are the first positive and negative deflections, respectively. N250 is functionally defined as having a greater negative deflection for target recognition (Tanaka et al., 2006).

subject, the value of each component for each trial in each condition was then determined as a sum of the seven data points surrounding and including the subject's component center (approximating a 13.6-ms integral about the component center).

This area measure was computed for each trial, rather than across the trial-averaged data, to facilitate modeling of the data using a general linear model (GLM). Though commonly employed in fMRI analysis, GLM is rarely applied to ERPs. However, the GLM approach is methodologically superior for studies of similarity space, as it provides unbiased parameter estimates of both the “direct effect” (Aguirre, 2007) of morph identity and of carry-over effects associated with similarity to the preceding face. If direct effects alone had been measured, the amplitude for, e.g., the extreme Face A would be influenced by the tendency of that extreme Face A to be preceded by dissimilar faces, and thus be subject to less adaptation. Simultaneous estimation of the direct and carry-over effects in the context of a counterbalanced stimulus order allows the estimates to be efficient and unbiased. Similarly, as each condition in the non-geometric bias model represented a different subset of face identities, the simultaneous modeling of this effect and the direct effects ensure unbiased estimation of each.

For each subject, the data for each component (P100, N170, P200, N250) were entered into a general linear model composed of 11 covariates. Five covariates coded for the particular morph identity (Figure 1A) presented on any one trial: the “direct effect” of a given morph identity upon the amplitude of an ERP component. The remaining covariates modeled carry-over effects, or the effect of the status of the prior trial upon response amplitude for a given trial. Five of these covariates modeled the different sizes of change in stimulus identity between one trial and the next ($\Delta 0\%$, $\Delta 25\%$, $\Delta 50\%$, $\Delta 75\%$, $\Delta 100\%$; Figure 4A); each covariate modeled those trials that had the given amount of identity change. A final covariate modeled asymmetric bias and was set to have a positive value for trials in which the preceding trial was at the extreme of the morph continuum (0% or 100%) and the current trial at the center (50%), and a negative value for transitions in the other direction (from 50% to 0% or 100%). Trials in which the target face was presented, and the trials that followed target face presentations, were excluded. The estimates obtained from this first-order analysis were then collected across subjects into a second-order, random effects ANOVA analysis to test hypotheses of interest.

Results

In this experiment, we explored the time course of perceptual similarity by recording ERPs during face perception. Given that behavioral judgment of similarity

has been hypothesized to consist of geometric effects of stimulus similarity and non-geometric effects of stimulus-specific bias, we tested if graded neural adaptation in the ERP data was consistent with this dual-effects model.

Behavioral measure of perceptual similarity

To confirm that the stimuli were linearly ordered in perceived similarity, we collected a behavioral measure of similarity in all subjects. All subjects participated in a paired-discrimination task using the face stimuli. Accuracy across subjects was sufficient (mean d' 2.15) to allow an analysis of reaction time effects. An MDS analysis was conducted for each subject on the average reciprocal reaction time for each face pairing, and then averaged across subjects. Figure 3 presents the position of the five faces on the first MDS dimension, which accounts for 55% of the variance. As can be seen, the first dimension contained a monotonic ordering of the stimuli, with somewhat greater spacing of the faces away from the 50% morph. There was substantial agreement across subjects on the perceptual similarity of the stimuli as demonstrated by the small across-subject error bars. This ordering of the stimuli confirms that, as expected from the stimulus design, subjects perceived a monotonic perceptual change in identity across the face morph continuum.

Geometric effect of stimulus similarity in ERP responses

ERP data were collected while subjects viewed a continuous stream of stimuli from the face continuum, presented in a counterbalanced order. ERP responses were assessed in relation to the identity of the face being presented, as well as the relationship of the current stimulus to the prior stimulus.

We first tested for a geometric effect of stimulus similarity based on the absolute metric distance from the

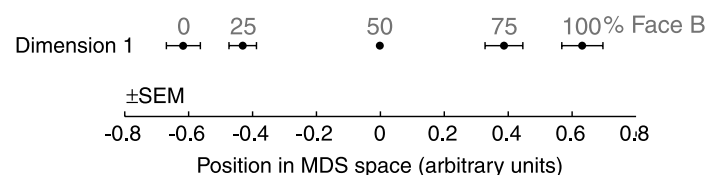
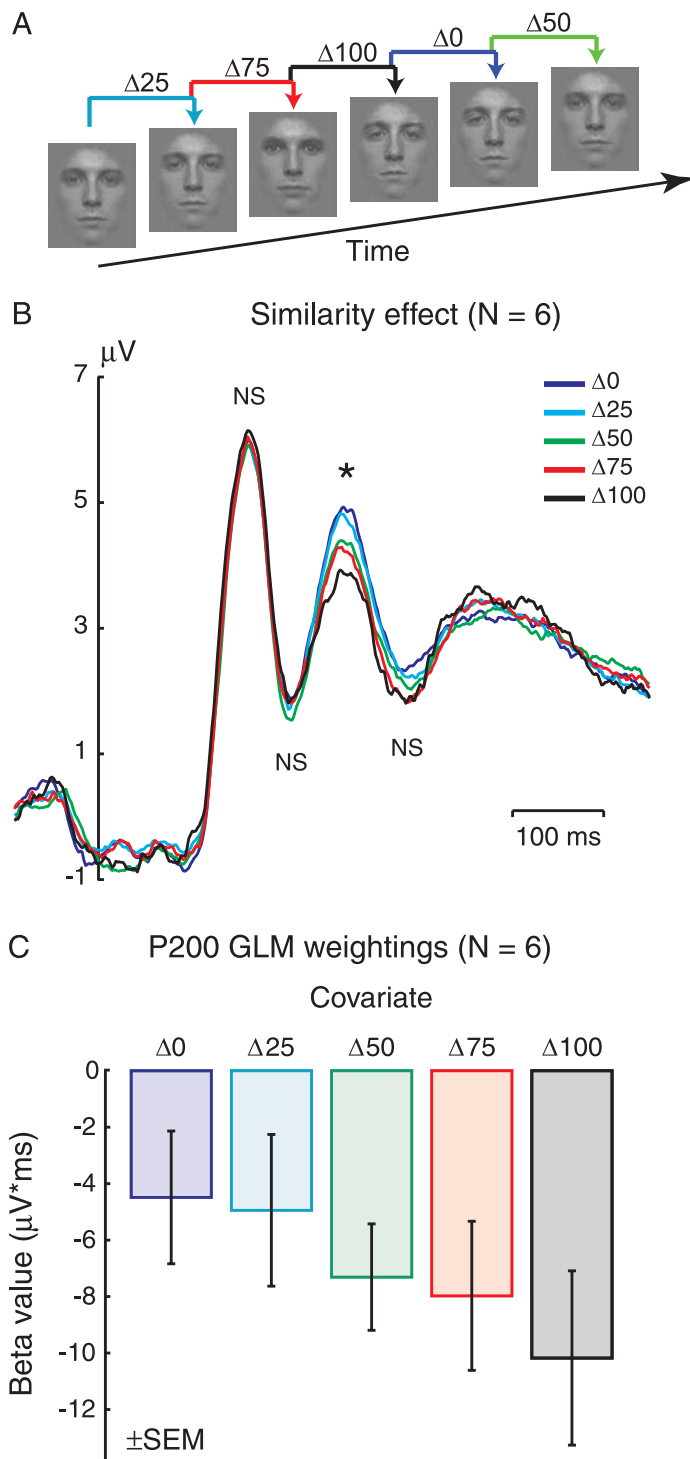


Figure 3. Behavioral results. Inverse reaction times from a paired discrimination task from each of six subjects were entered into a multi-dimensional scaling analysis, with the resulting coordinates centered about the 50% face. The first dimension of the resulting model is displayed, which orders the faces monotonically along the morph continuum. This first dimension accounts for 55% of the variance. Error bars indicate plus/minus standard error of the mean across subjects.

preceding stimulus to the current stimulus along the face identity continuum. Data from each trial were binned depending on the morph distance between the face shown and the previous image, resulting in five similarity distances ($\Delta 0$, $\Delta 25$, $\Delta 50$, $\Delta 75$, $\Delta 100$). Thus, a distance of $\Delta 0$ would be a repetition of the identical stimulus, whereas $\Delta 100$ represented a stimulus at one extreme of the morph continuum following the face at the opposite extreme (Figure 4A).



Because of the monotonic ordering of the perceptual similarity space used here, we would predict that the representation of metric stimulus similarity should change monotonically as a function of perceptual distance. In particular, given previous findings of neural adaptation in MEG (Furl et al., 2007; Harris & Nakayama, 2007, 2008) and ERP (Itier & Taylor, 2002; Jacques & Rossion, 2006; Kovács et al., 2006), we would predict greatest attenuation for $\Delta 0$, the identical repetition condition, with decreasing adaptation for increasing perceptual distances between stimuli.

Grand average waveforms across all significant ERP channels (Figure 2) for each perceptual distance condition are displayed in Figure 4B. While the early perceptual P100 and N170 components showed no discernible effect of stimulus similarity, a graded adaptation effect is clearly visible between the N170 and N250 components. The most positive deflection for this component occurs in the $\Delta 0$ condition, with decreasing amplitudes for greater perceptual distances. Modulation of the P200 component, therefore, appears to index the earliest stage of processing associated with computations of metric stimulus similarity. Caution is required in interpreting these average plots, however. As discussed previously, apparent graded responses in the waveforms could result not from an adaptation effect, but instead from the unbalanced representation of particular face identities in a given dissimilarity pair (see Supplementary Table 1).

To test this finding in an unbiased manner, beta values from the general linear model were obtained for each subject and component, representing the weight on covariates modeling each absolute distance condition. These measures are independent of any "direct effect" of stimulus identity (e.g., a hypothetically larger response to the extreme Face A or Face B). A repeated-measures ANOVA with component (P100, N170, P200, N250), and perceptual distance ($\Delta 0$, $\Delta 25$, $\Delta 50$, $\Delta 75$, $\Delta 100$) as factors showed a significant interaction between component and distance [$F(12, 60) = 5.05$, $p = 0.00001$], confirming that the effect of stimulus similarity is not seen for all components. Follow-up one-way repeated-measures ANOVAs for each

Figure 4. Geometric effect of similarity. (A) Trials were grouped based upon the metric distance of the preceding stimulus to the current stimulus along the morphed face continuum. Trials in which the target face was the current or preceding stimulus were excluded from analysis. (B) Grand average waveforms (across all significant sensors; Figure 2) comparing each distance transition condition. A significant interaction of component and distance condition was observed, and within the P200 component, there was a significant effect of distance (asterisk). Y-axis is aligned to stimulus onset. (C) Group average beta values from P200 for the five covariates modeling each distance condition in the general linear model. A significant effect of distance was observed, with a significant linear contrast. Error bars correspond to the between-subject SEM.

component found a significant main effect of distance for the P200 [$F(4, 20) = 6.01, p = 0.002$] (Figure 4C), but no other components (all F -tests $< 2.8, ps > 0.05$). The adaptation effect at the P200 was well modeled by a linear contrast [$F(1, 5) = 12.9, p = 0.016$]. While a similar ordering of the adapted response is visible in the grand average waveform at the later N250 (Figure 4B), this effect was not significant ($F(4, 20) = 3.38, p = 0.125$).

Therefore, these data suggest that neural sensitivity to perceptual similarity begins within the first 400 ms of

perceptual processing after stimulus onset. While the early perceptual P100 and N170 components do not show an effect of stimulus similarity, graded neural adaptation related to symmetric perceptual distance can be seen at the stage of processing following N170, the P200 response. Along with its temporal position between N170 and N250, this finding could be interpreted as placing the P200 at an intermediate cognitive stage between perceptual and mnemonic encoding.

Non-geometric effect of asymmetric bias in ERP responses

In addition to the geometric representation of stimulus similarity, we also tested for non-geometric, asymmetric neural representation of the stimulus space. Given behavioral findings demonstrating a bias for more “prototypical” stimuli (Op de Beeck, Wagemans, & Vogels, 2003), we hypothesized that the central face in the set, being an average of the faces at the extremes, would yield a differential effect on neural adaptation depending on whether it was a prior or current stimulus.

We compared the response on trials in which the central face is preceded by either of the two faces on the extreme of the stimulus space to trials in which the extreme faces are preceded by the central face (Figure 5A). Crucially, both of these conditions represent the same metric distance transition ($\Delta 50$) but vary in the direction of transition (“toward the center” of the stimulus space, and “toward the extremes”). Previous work has proposed that extreme stimuli preceded by more central or prototypical stimuli are perceived as more dissimilar than central stimuli preceded by extremes (Op de Beeck et al., 2003; Tversky, 1977). Therefore, we predicted that neural adaptation would be sensitive to the direction of stimulus transition, with greater neural adaptation for transitions toward the center and less adaptation toward the extreme.

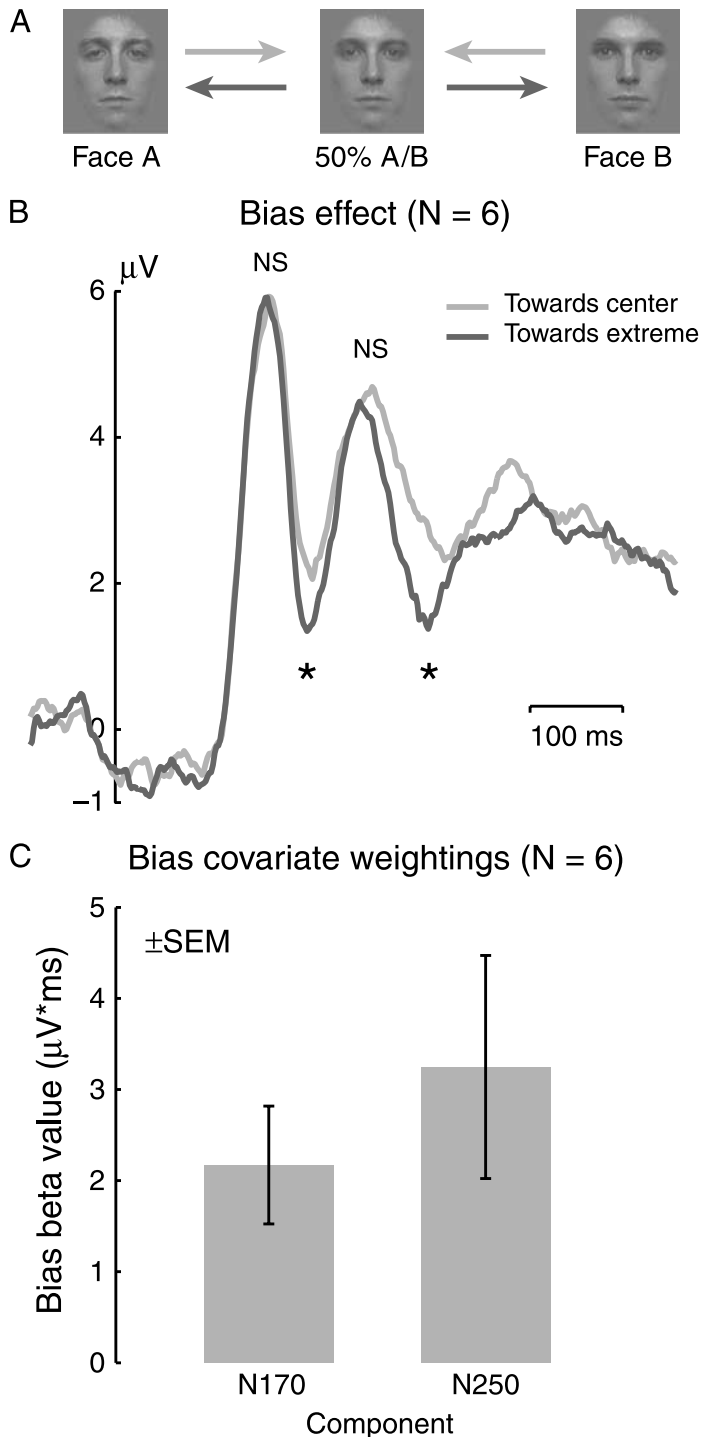
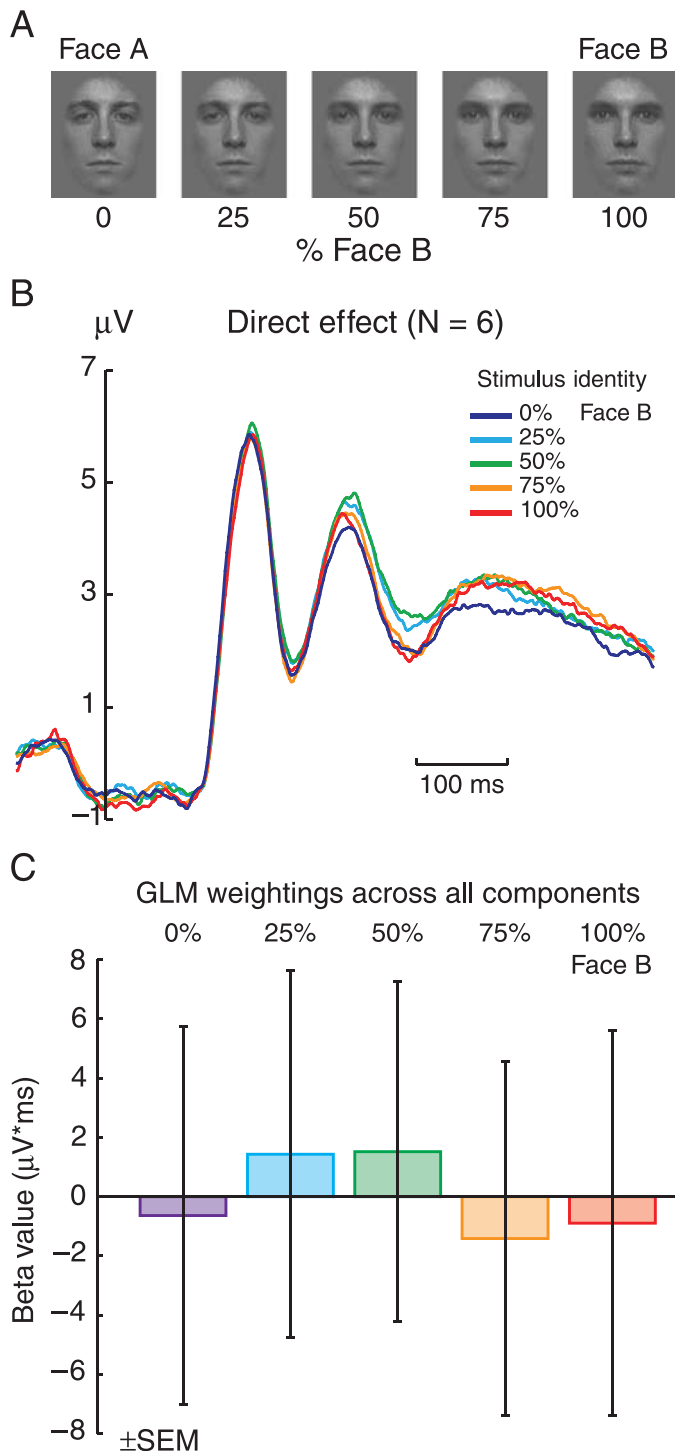


Figure 5. Non-geometric effects of similarity. (A) Trials were grouped based upon the direction of transition. “Toward center” trials were those in which the 50% face was presented following a face at either extreme of the morph continuum. “Toward extreme” trials had the opposite transition. (B) Grand average waveforms (across all significant sensors; Figure 2) comparing each condition. A significant interaction of component and direction condition was observed, and significant effects of direction were observed within the N170 and N250 components (asterisks). Y-axis is aligned to stimulus onset. (C) Group average beta values from the N170 and N250 components for the covariate modeling the bias condition. Both components had significant weighting on the bias covariate, showing greater adaptation for the “toward center” transition in line with described prototype effects. These beta estimates are corrected for any “direct” effect of stimulus identity (i.e., a tendency for a larger amplitude response to “extreme” faces).

A group average of the two bias conditions is plotted in [Figure 5B](#). In line with our predictions, transitions from the center of the stimulus space toward the extremes yield a greater negative deflection—but only at the N170 and N250 components. In contrast, P100 and P200 display equal adaptation for both presentation orders. Again, these average waveforms confound direct and carry-over effects due to unbalanced representation of transitions and face identities (see [Supplementary Table 2](#)).



To evaluate the statistical significance of this effect, we modeled the stimulus transition as a covariate in a general linear model analysis. Loading on this covariate indexes the asymmetric carry-over effect of the transition, independent of other symmetric carry-over or direct effects. A repeated-measures ANOVA for the single bias covariate with component (P100, N170, P200, N250) as a factor showed a significant main effect of component [$F(3, 15) = 7.536$, $p = 0.003$]. Follow-up one-sample t -tests across subjects indicated that this asymmetric bias is significant in the N170 [$t(5) = 3.36$, $p = 0.02$] and N250 [$t(5) = 2.65$, $p = 0.045$] components ([Figure 5C](#)).

Thus, asymmetric bias effects also occur within the first several hundred milliseconds of visual processing. Interestingly, in contrast to the N170 and N250 responses, P200 showed no significant asymmetric bias. This suggests, regardless of how geometric and non-geometric effects of similarity interact psychologically, that the earliest neural stages associated with these computations are temporally separated. The visible asymmetric bias at the relatively early N170 response may be indicative that such bias effects need not rely on higher level conceptual processing but may be extracted relatively rapidly and early in the visual processing stream.

Direct effects of stimulus identity on ERP responses

Finally, we examined the “direct” effect of stimulus identity upon the ERP response. Studies of “prototype” responses in fMRI to faces, for example, have reported that there is a larger amplitude of neural response to distinct, as opposed to typical, stimuli (Loffler et al., 2005).

A group average of the stimulus identity conditions is presented in [Figure 6B](#). Some separation between the identities is visible in the P200 and N250 components, perhaps consistent with a differential response to the extreme stimuli from the morph continuum as compared to the center. As discussed previously, however, these effects may be confounded by carry-over effects. For instance, a grand average waveform for the “direct” effect

Figure 6. Direct effects of stimulus identity. (A) Trials were grouped based upon the identity along the morph continuum shown. (B) Grand average waveforms (across all significant sensors; [Figure 2](#)) comparing each identity condition. A significant main effect of identity was observed, but no significant interaction of identity and component. Y-axis is aligned to stimulus onset. (C) Group average beta values collapsed across component are shown. As there was no significant main effect of component, or interaction of component with identity condition, beta values were mean-centered within component for each subject, averaged across component for each subject, and then averaged across subject for display. Error bars correspond to between-subject SEM of mean-centered, across-component averages.

of the 50% morph face is confounded by the fact that the 50% morph is, on average, more often preceded by similar faces by virtue of its central location; and thus more subject to adaptation. Similarly, a postulated differential response to the 50% morph face compared to the extreme faces (a “direct” effect) might confound the non-geometric bias effects without concurrent modeling.

To examine direct effects in an unbiased manner, we obtained the beta values associated with the amplitude of the ERP response to each face identity, after accounting for the adaptation and bias effects. A repeated-measures ANOVA was then performed with each identity covariate (0%, 25%, 50%, 75%, 100% Face B identities) and component (P100, N170, P200, N250) as factors. A significant main effect of identity was found [$F(4, 20) = 5.444$, $p = 0.004$], but the interaction of identity and component was nonsignificant [$F(12, 60) = 1.400$, $p = 0.191$], suggesting that this main effect of identity did not differentially modulate any component in particular. Figure 6C presents the average across subjects and components of the response to each face identity. The pattern of responses does not correspond readily to a simple model of prototype or geometric effects.

Discussion

Psychological models of perceptual proximity, the subjective judgment of “likeness” between stimuli, have historically drawn a distinction between two factors or processes: representation of simple metric distance between stimuli and stimulus-specific bias. Quantified in models such as the “additive similarity and bias” model (Holman, 1979; Nosofsky, 1991), this two-part framework separating geometric and non-geometric effects has guided our understanding of how the visual system represents variation between stimuli.

What are the neural correlates of these processes? We examined this question using a continuous carry-over design (Aguirre, 2007) in ERP. Previously used in fMRI, continuous carry-over designs allow measurement of graded neural adaptation, and therefore better characterization of the neural representation of perceptual similarity space. Using this paradigm with a set of ordered, morphed faces in ERP, we tested a dual-effects model of perceptual similarity against a temporal framework of early visual evoked components previously associated with face processing.

Modeling transitions between stimulus presentations in terms of absolute metric distance along our morphed face continuum, we found graded neural adaptation consistent with metric stimulus similarity at a component between the N170 and N250 responses. Modulation of P200 was related to perceptual similarity, with greater positive deflection for smaller perceptual distances (Figure 4).

The temporal position of this component suggests that computation of metric stimulus similarity begins within the first several hundred milliseconds of stimulus presentation, although after the earliest stages of perceptual processing indexed by the P100 and N170 components. Adaptation of a neuroimaging signal that is proportional to stimulus similarity can result from a cortical region that codes stimulus identity by a population code (Aguirre, 2007; Drucker, Kerr, & Aguirre, 2009). This suggests that, at the P200 stage, a neural population code for facial identity is evoked that reflects geometric effects of similarity. It is also possible that another neural mechanism apart from adaptation (e.g., a re-entrant masking effect; Kotsoni, Csibra, Mareschal, & Johnson, 2007) is responsible for this parametric modulation. In either case, these data are among the first to place a neural signature of geometric similarity coding within a definite time window, arising as early as 200 ms after stimulus presentation.

We also modeled the effects of asymmetric bias (Op de Beeck et al., 2003; Tversky, 1977). Neural markers of such a non-geometric similarity effect were found for the N170 and N250 components (Figure 5). While both the N170 and N250 components show sensitivity to asymmetric transitions positioned about the center of the stimulus space, the P200 does not. Thus, not only have we found neural correlates of perceptual proximity processing within relatively early stages of perceptual processing, but we also demonstrate that the encoding of metric stimulus similarity and asymmetric bias are temporally distinct.

Our model of non-geometric similarity effects is based upon the notion of a “prototype” effect (Op de Beeck et al., 2003; Tversky, 1977). Two stimuli are perceived as more proximal when the more prototypical or average stimulus is presented following another one less so, and less proximal in the reverse case. There are other non-geometric bias effects that might be considered. In studies of magnitude estimation, for example, the response to a stimulus tends to be larger when the preceding stimulus intensity was greater. This “assimilation” effect is commonly seen for stimuli in which one end of the continuum is “larger” (DeCarlo & Cross, 1990). The opposite, “contrast” effect is also observed. A model for this directional bias in neuroimaging data is considered in Aguirre (2007) and is orthogonal to the “prototype” effect just discussed. While the “prototype” model of bias is symmetric about the center of the stimulus space, directional bias is inversely symmetric toward each extreme. Directional bias has been observed in perceptual adaptation effects for face identity (Leopold, O’Toole, Vetter, & Blanz, 2001), gender (Webster, Kaping, Mizokami, & Duhamel, 2004), and attractiveness (Rhodes, Jeffery, Watson, Clifford, & Nakayama, 2003). We tested for directional bias effects in our ERP study but found no significant effect (data not shown). This is not surprising as our stimuli were a morph between two faces of equal distinctiveness, as opposed to the stimuli of, e.g., Leopold

et al. (2001) in which one end of the continuum was a distinctive face and the other a prototypical or average face.

A perceptual prototype may arise from long-term exposure to stimuli of a given class, from the local context of a set of stimuli in an experiment, or both. Our study did not distinguish between these two types of prototype. The center point of our stimulus continuum may have achieved prototype status as it was a more “average” face in general, or because it was the central tendency of this particular stimulus set. These possibilities might be distinguished through the use of an unbalanced face continuum in which the “middle” face in the local context of the experimental set is not the most average at the global level.

Related to this point, it is worth noting that while we observed neural prototype effects for both the N170 and N250 components, it is possible that these distinct components are related to different prototype effects. For the N170 in particular, we might expect that the “prototype” effect reflects a local stimulus effect, driven by the experimental stimulus space alone. Previous work has demonstrated a lack of adaptation in N170 to within-class features of faces, including eye-gaze direction (Schweinberger, Kloth, & Jenkins, 2007) and gender (Kloth, Schweinberger, & Kovács, 2009). These findings suggest that N170 adapts in a broad categorical fashion to faces and not to within-category features, such as global face distinctiveness. Taken together with the apparent role of N170 in structural encoding (Bentin et al., 1996; Rossion et al., 2000), we would suggest that the “prototype” effect observed in N170 might reflect a rapid, implicit extraction of local central tendency (i.e., within the experimental stimulus space). In contrast, as N250 is thought to reflect access to stored face representations (Tanaka et al., 2006), it is possible that the non-geometric effect observed in this component indexes transitions about a stored, global face “prototype”. While these interpretations rely on the characteristics of the underlying components, future experiments that dissociate local and global face prototypes in the manner described above could characterize putatively separable non-geometric similarity effects in a component-independent manner.

Finally, a notable methodological feature of this study was the concurrent measurement and separation of the direct effects of each stimulus from carry-over effects of adaptation and asymmetric bias. Without explicit modeling, these effects are confounded, rendering it unclear whether effects reflect perceptual proximity per se, or a combination of adaptation and identity effects. This potential confound exists in several studies of face representation. For example, Loffler et al. (2005) used a block design in fMRI to demonstrate increasing BOLD signal in the fusiform face area (FFA) in response to groups of faces of increasing “distinctiveness”. The authors define “distinctiveness” as distance along putatively orthogonal identity axes extending from a central “mean” face. This design focuses primarily on non-geometric prototype and

identity effects. However, their observed decrease in BOLD signal for face blocks more proximal to the mean could represent neural adaptation indexing geometric effects of metric distance, or some combination of geometric and non-geometric effects.

Likewise, in an fMRI study using a similar facial identity morph continuum to ours, Jiang et al. (2006) reported non-linear BOLD adaptation in response to increasing metric distance. The authors interpreted this finding as suggesting that neural adaptation would asymptote for greater metric stimulus distances, something we do not observe in our data. In their experimental design, Jiang et al. (2006) use a traditional paired-presentation paradigm with the adapting stimuli only located at the extreme of the morph continuum and test stimuli at $\Delta 30$, $\Delta 60$ and $\Delta 90$ metric distances. It is possible with this design that the unbalanced frequency of stimulus presentation introduces a non-geometric similarity effect such as the “relative prominence” bias presented by Johansson (2000), or an asymmetry driven by exposure frequency as presented by Polk et al. (2002). Thus while Jiang et al. (2006) suggest their data reflects non-linear (asymptotic) encoding of metric linear distance, our findings suggest that their data could reflect a combination of geometric effects and non-geometric effects.

Conclusions

Our results provide evidence for the dissociation in neural coding of non-geometric “prototype” effects from the geometric effects of stimulus similarity, supporting psychological models of the two elements as separate factors in the perception of proximity. Using a continuous carry-over design in ERP, in conjunction with a principled GLM approach to distinguish geometric and non-geometric processing, we find that these different effects occur at discrete temporal stages of face processing. These findings should expand our understanding of neural similarity, offer new avenues for exploring global and local prototype effects, and encourage more careful consideration of the complexity of stimulus space representations in the brain.

Acknowledgments

This work was supported by K08 MH 7 2926-01 and a Burroughs-Wellcome Career Development Award. Development of the MacBrain Face Stimulus Set was overseen by Nim Tottenham and supported by the John D. and Catherine T. MacArthur Foundation Research Network on Early Experience and Brain Development. Please contact

Nim Tottenham at tott0006@tc.umn.edu for more information concerning the stimulus set.

Commercial relationships: none.

Corresponding author: Geoffrey Karl Aguirre.

Email: aguirreg@mail.med.upenn.edu.

Address: Department of Neurology, University of Pennsylvania, 3 West Gates, 3400 Spruce Street, Philadelphia, PA 19104, USA.

References

- Aguirre, G. K. (2007). Continuous carry-over designs for fMRI. *Neuroimage*, *35*, 1480–1494.
- Bentin, S., Allison, T., Puce, A., & Perez, E. (1996). Electrophysiological studies of face perception in humans. *Journal of Cognitive Neuroscience*, *8*, 551–565.
- Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, *10*, 433–436.
- DeCarlo, L. T., & Cross, D. V. (1990). Sequential effects in magnitude scaling: Models and theory. *Journal of Experimental Psychology: General*, *119*, 375–396.
- Delorme, A., & Makeig, S. (2004). EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, *134*, 9–21.
- Di Russo, F., Martínez, A., Sereno, M. I., Pitzalis, S., & Hillyard, S. A. (2001). The cortical sources of the early components of the visual evoked potential. *Human Brain Mapping*, *15*, 95–111.
- Drucker, D., & Aguirre, G. (2009). Different spatial scales of shape similarity representation in lateral and ventral LOC. *Cerebral Cortex*, *19*, 2269–2280.
- Drucker, D. M., Kerr, W. T., & Aguirre, G. K. (2009). Distinguishing conjoint and independent neural tuning for stimulus features with fMRI adaptation. *Journal of Neurophysiology*, *101*, 3310–3324.
- Furl, N., van Rijsbergen, N. J., Treves, A., Friston, K. J., & Dolan, R. J. (2007). Experience-dependent coding of facial expression in superior temporal sulcus. *Proceedings of the National Academy of Sciences of the United States of America*, *104*, 13485–13489.
- Grill-Spector, K., & Malach, R. (2001). fMR-adaptation: A tool for studying the functional properties of human cortical neurons. *Acta Psychologica*, *107*, 293–321.
- Harris, A., & Nakayama, K. (2007). Rapid face-selective adaptation of an early extrastriate component in MEG. *Cerebral Cortex*, *17*, 63–70.
- Harris, A., & Nakayama, K. (2008). Rapid adaptation of the M170 response: Importance of face parts. *Cerebral Cortex*, *18*, 467–476.
- Henson, R. N. A., & Rugg, M. D. (2003). Neural response suppression, haemodynamic repetition effects, and behavioural priming. *Neuropsychologia*, *41*, 263–270.
- Holman, E. W. (1979). Monotonic models for asymmetric proximities. *Journal of Mathematical Psychology*, *20*, 1–15.
- Itier, R. J., & Taylor, M. J. (2002). Inversion and contrast polarity reversal affect both encoding and recognition processes of unfamiliar faces: A repetition study using ERPs. *Neuroimage*, *15*, 353–372.
- Itier, R. J., & Taylor, M. J. (2004). N170 or N1? Spatiotemporal differences between object and face processing using ERPs. *Cerebral Cortex*, *14*, 132–142.
- Jacques, C., & Rossion, B. (2006). The speed of individual face categorization. *Psychological Science*, *17*, 485–492.
- Jiang, X., Rosen, E., Zeffiro, T., Vanmeter, J., Blanz, V., & Riesenhuber, M. (2006). Evaluation of a shape-based model of human face discrimination using fMRI and behavioral techniques. *Neuron*, *50*, 159–172.
- Johannesson, M. (2000). Modelling asymmetric similarity with prominence. *British Journal of Mathematical and Statistical Psychology*, *53*, 121–139.
- Kloth, N., Schweinberger, S., & Kovács, G. (2009). Neural correlates of generic versus gender-specific face adaptation. *Journal of Cognitive Neuroscience*, *22*, 2345–2356.
- Kotsoni, E., Csibra, G., Mareschal, D., & Johnson, M. H. (2007). Electrophysiological correlates of common-onset visual masking. *Neuropsychologia*, *45*, 2285–2293.
- Kovács, G., Zimmer, M., Bankó, E., Harza, I., Antal, A., & Vidnyánszky, Z. (2006). Electrophysiological correlates of visual adaptation to faces and body parts in humans. *Cerebral Cortex*, *16*, 742–753.
- Krumhansl, C. L. (1978). Concerning the applicability of geometric models to similarity data: The interrelationship between similarity and spatial density. *Psychological Review*, *85*, 445–463.
- Kruskal, J. B., & Wish, M. (1978). *Multidimensional scaling quantitative applications in the social sciences*. Thousand Oaks, CA: Sage.
- Leopold, D. A., O'Toole, A. J., Vetter, T., & Blanz, V. (2001). Prototype-referenced shape encoding revealed by high-level aftereffects. *Nature Neuroscience*, *4*, 89–94.
- Liu, J., Harris, A., & Kanwisher, N. (2002). Stages of processing in face perception: An MEG study. *Nature Neuroscience*, *5*, 910–916.
- Liu, J., Higuchi, M., Marantz, A., & Kanwisher, N. (2000). The selectivity of the occipitotemporal M170 for faces. *NeuroReport*, *11*, 337–341.

- Loffler, G., Yourganov, G., Wilkinson, F., & Wilson, H. R. (2005). fMRI evidence for the neural representation of faces. *Nature Neuroscience*, 8, 1386–1390.
- Nosofsky, R. (1991). Stimulus bias, asymmetric similarity, and classification. *Cognitive Psychology*, 23, 94–140.
- Op de Beeck, H., Wagemans, J., & Vogels, R. (2001). Inferotemporal neurons represent low-dimensional configurations of parameterized shapes. *Nature Neuroscience*, 4, 1244–1252.
- Op de Beeck, H., Wagemans, J., & Vogels, R. (2003). Asymmetries in stimulus comparisons by monkey and man. *Current Biology*, 13, 1803–1808.
- Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, 10, 437–442.
- Polk, T. A., Behensky, C., Gonzalez, R., & Smith, E. E. (2002). Rating the similarity of simple perceptual stimuli: Asymmetries induced by manipulating exposure frequency. *Cognition*, 82, B75–B88.
- Renninger, L. W., & Malik, J. (2004). When is scene identification just texture recognition? *Vision Research*, 44, 2301–2311.
- Rhodes, G., Jeffery, L., Watson, T. L., Clifford, C. W. G., & Nakayama, K. (2003). Fitting the mind to the world: Face adaptation and attractiveness aftereffects. *Psychological Science*, 14, 558–566.
- Rossion, B., Gauthier, I., Tarr, M. J., Despland, P., Bruyer, R., Linotte, S., et al. (2000). The N170 occipito-temporal component is delayed and enhanced to inverted faces but not to inverted objects: An electrophysiological account of face-specific processes in the human brain. *NeuroReport*, 11, 69–74.
- Sams, M., Hietanen, J. K., Hari, R., Ilmoniemi, R. J., & Lounasmaa, O. V. (1997). Face-specific responses from the human inferior occipito-temporal cortex. *Neuroscience*, 77, 49–55.
- Schweinberger, S. R., Kloth, N., & Jenkins, R. (2007). Are you looking at me? Neural correlates of gaze adaptation. *NeuroReport*, 18, 693–696.
- Schweinberger, S. R., Pickering, E. C., Jentzsch, I., Burton, A. M., & Kaufmann, J. M. (2002). Event-related brain potential evidence for a response of inferior temporal cortex to familiar face repetitions. *Cognitive Brain Research*, 14, 398–409.
- Shepard, R. (1964). Attention and the metric structure of the stimulus space. *Journal of Mathematical Psychology*, 1, 54–87.
- Tanaka, J. W., Curran, T., Porterfield, A. L., & Collins, D. (2006). Activation of preexisting and acquired face representations: The N250 event-related potential as an index of face familiarity. *Journal of Cognitive Neuroscience*, 18, 1488–1497.
- Torgerson, W. S. (1965). Multidimensional scaling of similarity. *Psychometrika*, 30, 379–393.
- Tottenham, N., Tanaka, J. W., Leon, A. C., McCarry, T., Nurse, M., Hare, T. A., et al. (2009). The NimStim set of facial expressions: Judgments from untrained research participants. *Psychiatry Research*, 168, 242–249.
- Tversky, A. (1977). Features of similarity. *Psychological Review*, 84, 327–352.
- Webster, M. A., Kaping, D., Mizokami, Y., & Duhamel, P. (2004). Adaptation to natural facial categories. *Nature*, 428, 557–561.