The effects of parts, wholes, and familiarity on face-selective responses in MEG

Alison M. Harris
Department of Neurology, University of Pennsylvania, Philadelphia, PA, USA
Geoffrey K. Aguirre
Department of Neurology, University of Pennsylvania, Philadelphia, PA, USA

Although face perception is commonly characterized as holistic, as opposed to part-based, we have recently shown that both face parts and wholes are represented in “face-selective” cortical regions, with greater adaptation of holistic representations for familiar faces (A. Harris & G. K. Aguirre, 2008). Here we investigate the time course of these holistic and part-based face processing effects using magnetoencephalography (MEG). We examined “face-selective” components at early (~170–200 ms) and later (~250–450 ms) latencies in occipitotemporal sensors. While both “M170” and “M400” components showed significantly larger responses for familiar versus unfamiliar faces, neither exhibited a main effect of holistic versus part-based processing. These data affirm the existence of part-based “face-selective” representations, and additionally demonstrate that such representations are present from relatively early stages of face processing. However, only the later M400 component showed a modulatory effect of familiarity similar to that previously seen with fMRI, with a larger response to familiar faces in the holistic condition. Likewise, behavioral recognition was significantly correlated with the M400, not the M170, and only in the holistic condition. Together, these data suggest that, while face parts are represented from the earliest stages of face perception, modulatory effects of familiarity occur later in the face processing stream.

Keywords: face perception, neurophysiology, M170, configural processing

Introduction

With its myriad cues to social status, the face is arguably one of the most important visual stimuli regularly encountered by human observers. Yet despite their relative homogeneity, we can readily parse faces for complex information such as identity, in some cases demonstrating high accuracy even across long periods of time (Bruck, Cavanagh, & Ceci, 1991).

In fact, a large behavioral literature suggests that face recognition is “special,” subserved by different mechanisms than the majority of visual stimuli. Key evidence in favor of this idea comes from the “face inversion effect” (Yin, 1969), the finding that inversion detrimentally affects perception of faces more than that of other stimuli (though see Husk, Bennett, & Sekuler, 2007).

Based on this result, it has been proposed that upright faces differ from other stimuli in that they are represented in a holistic or configural, as opposed to part-based, manner (Diamond & Carey, 1986; Farah, Wilson, Drain, & Tanaka, 1998): that is, in terms of the relationship between parts rather than the parts themselves. According to this model, the “face inversion effect” arises from the disruption of such holistic processing, perhaps in conjunction with a switch to a part-based object recognition system (Bartlett & Searcy, 1993; Farah, 1990; Searcy & Bartlett, 1996). This idea has received additional support from other behavioral data, including “composite” face recognition (Young, Hellawell, & Hay, 1987) and the “whole-versus-part superiority effect” (Tanaka & Farah, 1993), in which recognition accuracy for individual face parts is increased by presentation in the context of a whole face, versus in isolation.

Yet, despite the emphasis on holistic face processing in the literature, there is some evidence for the representation of face parts as well (Cabeza & Kato, 2000; Leder & Bruce, 1998; Macho & Leder, 1998). Cabeza and Kato (2000), for example, reported a “prototype effect” for face parts, that is, an increased likelihood of falsely recognizing a novel face composed of parts from previously-seen faces. Nor are part-based representations necessarily an exclusive function of the object recognition system: CK, a patient with severe object agnosia, can still identify isolated face features (Moscovitch, Winocur, & Behrmann, 1997).

From these data, it would appear that both wholes and parts are represented in the face processing stream. Electrophysiological recordings support this idea, finding some cells that respond selectively to whole faces (Bruce, Desimone, & Gross, 1981) and others sensitive to face parts (Perrett, Rolls, & Caan, 1982). Yet researchers using
functional magnetic resonance imaging (fMRI) have only recently examined the relation of “face-selective” regions of cortex to holistic and part-based processing (Schiltz & Rossion, 2006; Yovel & Kanwisher, 2005). Using composite and inverted faces, respectively, these studies found responses matching behavioral indices of holistic processing in areas of the fusiform and occipital gyri previously described as “face-selective” (Gauthier et al., 2000; Kanwisher, McDermott, & Chun, 1997; McCarthy, Puce, Gore, & Allison, 1997), particularly within the right middle fusiform gyrus (MFG; Schiltz & Rossion, 2006).

However, despite the implications of these studies for holistic processing, the issue of whether these face-selective regions also respond to parts has not been directly addressed. We have recently examined this question in fMRI (Harris & Aguirre, 2008) using a stereoscopic depth manipulation originally devised by Nakayama and colleagues (Nakayama, Shimojo, & Silverman, 1989). These stimuli use binocular disparity to create the percept of either a face behind a set of bars (Figure 1, left) or strips of a face floating in front of a background (Figure 1, right).

Note that, aside from the depth manipulation, these stimuli are identical in terms of their low-level stimulus properties. Yet, while the first case undergoes amodal completion and can be processed holistically, the latter cannot be completed and is therefore perceived in terms of its constituent parts. We have demonstrated this behaviorally using the “whole-versus-part superiority effect” (Tanaka & Farah, 1993), which we successfully replicated in the Back (holistic) but not the Front (part-based) depth condition (Harris & Aguirre, 2008). These results support the idea that the manipulation of stereoscopic depth can affect whether these stimuli undergo holistic or part-based processing.

Using these stimuli in fMRI, we have shown that face-selective areas of visual cortex respond to both Front and Back depth conditions. Additionally, inspired by behavioral data suggesting that familiar faces have more robust or reliable holistic representations (Buttle & Raymond, 2003; Young, Hay, McWeeny, Flude, & Ellis, 1985), we examined the effects of familiarity on holistic and part-based processing. Within the right middle fusiform region previously associated with holistic processing (Schiltz & Rossion, 2006), we found that there was greater adaptation in the Back (holistic) depth condition for familiar but not unfamiliar faces. Based on these results, we have argued that both face parts and wholes are represented in the face processing stream, and that the relative recruitment of these representations can be influenced by familiarity.

In the present experiment, we wished to extend our previous work by delineating the time course of holistic and part-based representations in the face processing stream. Despite its excellent capabilities for spatial localization, fMRI lacks the temporal resolution to distinguish neural events on a millisecond scale. Instead, we turned to magnetoencephalography (MEG), a neurophysiological technique argued to be particularly suited for measuring responses from ventral areas like the fusiform gyrus (Itier, Herdman, George, Cheyne, & Taylor, 2006). We focused on responses in two latency ranges: early (170–200 ms) and late (250–450 ms). For these two latency windows, we examined both the main effects of familiarity and processing type, as assessed by the depth manipulation, as well as the interaction of these factors.

In the early latency range, we quantified the response profile of the well-known M170 (in MEG) or N170 (in event-related potentials, or ERP), a “face-selective” component occurring approximately 170 ms after stimulus onset (Bentin, Allison, Puce, Perez, & McCarthy, 1996; Liu, Higuchi, Marantz, & Kanwisher, 2000; Sams, Hietanen, Hari, Ilmoniemi, & Lounasmaa, 1997). Although it is generally agreed that the M170/N170 reflects some aspect of perceptual encoding, the exact nature of this processing remains a point of contention.

Because it is consistently delayed and/or enhanced by face inversion (Bentin et al., 1996; Itier & Taylor, 2002; Rossion et al., 2000), the M170/N170 has been characterized as an index of configural processing (e.g., Rossion et al., 2000). Yet other evidence suggests that the M170/N170 responds primarily to the presence of face parts, particularly the eye region (Itier, Latinus, & Taylor, 2006; Schyns, Jentzsch, Johnson, Schweinberger, & Gosselin, 2003; Smith, Gosselin, & Schyns, 2004). In our own recent work, we have shown that rapid adaptation of the M170 response appears to be driven by the presence of face parts, not face configuration, supporting the idea that the processing indexed by this early response is part-based (Harris & Nakayama, 2008).

With regards to familiarity, results for the M170/N170 have been similarly inconsistent. While some studies have found no modulation of the M170/N170 response by familiarity (Bentin & Deouell, 2000; Eimer, 2000; Schweinberger, Pickering, Jentzsch, Burton, & Kaufmann,

![Figure 1. Illustration of stereoscopic depth manipulation derived from Nakayama et al. (1989). When the bars appear in front of the face (left), the face is amodally completed and holistic processing can occur. When the face appears in the frontal depth plane (right), the face cannot be completed and is perceived in terms of its parts.](http://jov.arvojournals.org/pdfaccess.ashx?url=/data/journals/jov/932850/ on 06/20/2017)
others have reported a larger response to familiar faces, particularly ones that are personally known (Caharel, Courtay, Bernard, Lalonde, & Rebai, 2005; Kloth et al., 2006). Likewise, experiments using trained or personally familiar faces degraded by phase-scrambled noise or masked short-duration presentation have also shown that M170 amplitude is correlated with recognition (Liu, Harris, & Kanwisher, 2002; Tanskanen, Näsänen, Ojanpää, & Hari, 2007), though these manipulations also affect the perceptibility of the stimulus.

In the later latency range of roughly 250 to 500 ms, various components have been described that are sensitive to face familiarity (Bentin & Deouell, 2000; Eimer, 2000; Tanskanen et al., 2007). These include effects related to stimulus repetition (Schweinberger et al., 2002), consolidation of familiar face representations (Tanaka, Curran, Porterfield, & Collins, 2006), and incongruency processing (Jemel, George, Olivares, Fiori, & Renault, 1999; Olivares, Iglesias, & Rodriguez-Holguin, 2003). Many of these later components are thought to reflect accessing of semantic information, as indicated by more central or frontal scalp distributions, though a mismatch negativity for purely visual stimuli has been reported at posterior temporal sensors (Olivares et al., 2003). In addition, there is evidence that some later components are associated with configural representations (Caharel, Fiori, Bernard, Lalonde, & Rebai, 2006).

However, since many of these responses are elicited only in the context of certain paradigms or task demands, in this experiment we chose not to explicitly search for a previously-defined component. Rather, we agnostically selected sensors in this later time window, here dubbed the “M400”, by searching for “face-selective” responses. Therefore, in the present study we examined the effects of processing type, as indexed by depth, and familiarity, as well as their interaction, for “face-selective” components at both early (170–200 ms) and late (250–450 ms) perceptual stages. Based on the existing literature, particularly our own results from a rapid adaptation paradigm (Harris & Nakayama, 2008), we predict that the M170 response will be indistinguishable for Front and Back depth conditions, as both contain face parts. In contrast, we expect effects of familiarity and the interaction of depth and familiarity to arise in the later M400 latency range, after the stage of processing indexed by the M170 response.

Methods

Subjects

14 individuals between the ages of 18 and 35 were recruited from local universities. All subjects had normal or contact-corrected vision and normal stereoscopic vision. 6 additional subjects were excluded due to excessive eye movement artifact (1 subject), improper head placement (1 subject), or failure to find “face-selective” (Face > House, described below) sensors in one or both hemispheres (4 subjects). Informed consent was obtained from all subjects, and procedures followed institutional guidelines and the Declaration of Helsinki.

Stimuli

Stimuli were created from 200 × 200 pixel grayscale photographs of famous celebrities, unfamiliar individuals, and houses (40 exemplars each). The unfamiliar faces were drawn from portfolio photographs of aspiring actors, and were selected to match the famous faces in gender, hair color and style, and attractiveness. These stimuli were placed into an image consisting of a gray background noise pattern overlaid with noise-patterned red and green stripes positioned to appear at 5 min or 9 min of disparity either in front of or behind the face stimulus when viewed with anaglyphic (red/green) glasses. The finished stimuli were 250 × 250 pixels, subtending 18.1° × 18.1° of visual angle, and were presented on a black background.

Procedure

The experiment consisted of 6 conditions: 3 stimulus categories (Famous, Unfamiliar, House) crossed with 2 depths (Back, Front). 80 trials per condition were presented randomly interleaved. Each stimulus was displayed for 500 ms with an inter-trial interval jittered between 1.1 and 1.4 s, during which a fixation image of the gray noise pattern without stripes was displayed (Figure 2). Subjects were instructed to monitor for the appearance of flower stimuli (40 exemplars, 2 depths), which comprised 14.3% of the total trials. Target trials...
were randomly intermixed with the experimental trials, but were excluded from analysis.

In order to obtain a canonical M170 for comparison, after the main experiment subjects completed a short “localizer” experiment with unoccluded faces, houses, and everyday objects (100 exemplars each), randomly interleaved. Stimuli were presented on a white background with a black fixation cross (14.6° × 14.6° visual angle) for 300 ms (ITI jittered between 900 and 1100 ms); subjects were instructed to remove the red/green anaglyphic glasses and passively view these stimuli.

13 of 14 subjects also performed a familiar/unfamiliar judgment task with a subset of the face stimuli to assess their familiarity with the celebrities depicted in the Famous condition.

MEG data analysis

Data recordings were made using a 275-channel whole-head system (one sensor excluded) with SQUID-based third-order gradiometer sensors (CTF, VSM MedTech) at the Children’s Hospital of Philadelphia. Magnetic brain activity was digitized in 1000 ms epochs (100 ms pre-, 900 ms post-stimulus onset) at a sampling rate of 600 Hz.

Data analysis was performed in MATLAB (Mathworks, Andover, MA) using the EEGLAB open source toolbox (Delorme & Makie, 2004). 700 ms epochs (100 ms pre-, 600 ms post-stimulus onset) for each condition in each subject were examined individually for artifacts (e.g., eye blinks), and up to 10 artifactual trials per condition (12.5%) were removed. Average waveforms were computed within the 700 ms window and low-pass filtered below 40 Hz.

Sensors were selected for analysis using a “sensor of interest” (SOI) approach (Liu et al., 2002), via a point-to-point t test comparing the Famous Back (face) and House Back (house) conditions (Figure 3). While early visual evoked responses such as the M170 can be selected by amplitude and latency alone (e.g., Harris & Nakayama, 2007), the use of the face versus house contrast ensured that the later “M400” component reflected perceptual processing rather than attention or task demands. In keeping with common practice, sensors selected by this comparison were labeled “face-selective,” though the term “preferential” has also been advocated (Pernet, Schyns, & Demonet, 2007).

Although late “face-selective” components have often been reported more frontally or centrally-distributed relative to the occipitotemporal M170/N170 response (Bentin & Deouell, 2000; Eimer, 2000; Schweinberger et al., 2002), preliminary analyses of the current data failed to show any clear “face-selective” pattern at central sensors. This result probably stems from methodological differences, such as the task demands or the comparison used for sensor selection. (Note that previous work describing a more central N400 for famous versus unfamiliar faces (Bentin & Deouell, 2000; Eimer, 2000) used a direct comparison of these two conditions rather than faces versus houses.)

Instead, preliminary examination of our data revealed that a subset of occipitotemporal sensors also showed a much greater response to faces than houses at later latencies. Therefore, only those sensors showing a significantly greater (t = 1.67) response to faces for a window of 20 time points in both the early (~200 ms) and late (~280–380 ms) range were used for analysis. The overlap in SOIs between subjects is shown in the scalp map at left in Figure 3.

Peak amplitude and latency of the M170 at designated SOIs were determined individually for each subject in each condition between 190 and 250 ms post-stimulus onset. For the later component (the “M400”), as there is often no clear peak in the data, we instead calculated the area under the curve (AUC) between 283 and 483 ms post-stimulus onset (Figure 3, gray box). These latency boundaries were chosen on the basis of the shape of the grand average waveform for the FamBack condition, but were applied to each subject’s data in each condition to obtain individual measurements of the M400. In a follow-up AUC analysis of the M170, the AUC was computed for a 100-ms window around the peak latency (50 ms before and 50 ms after) for each condition in each subject.

Data from the short “localizer” experiment were analyzed in the manner described above, but with a 500 ms time window (100 ms pre-, 400 ms post-stimulus onset). Subjects’ previously-defined SOIs were used to individually analyze their data, from which M170 peak amplitudes and latencies were determined for each condition.

Due to the nature of the magnetic field generated by electric currents in the brain, the B field corresponding to the M170 in the right hemisphere constitutes a magnetic
“sink,” which is commonly denoted by a negative sign; for averaged analyses, peak amplitudes in right hemisphere sensors were multiplied by $-1$ to correct for this polarity difference.

## Results

Average M170 peak amplitude and M400 AUC for each condition across 14 subjects are listed in Table 1. While the M170 and M400 responses are much larger for the Famous Back than House Back condition, as expected from our sensor selection procedure, all other face conditions elicit higher M170 and M400 responses than houses as well. Thus, consistent with our fMRI results, both part-based and holistic stimuli appear to evoke “face-selective” responses throughout the visual processing stream.

Figure 4 displays the grand average waveforms for familiar and unfamiliar faces as a function of Back (Figure 4A) and Front (Figure 4B) depth conditions. In the Back depth, associated with holistic processing, there is a clear effect of familiarity, with a larger response to familiar versus unfamiliar faces. Although this effect is largest in the M400 latency range, it is present even at the M170. In contrast, no such familiarity effect is visible for the part-based Front depth condition.

Separate repeated-measures ANOVAs on M170 peak amplitude and M400 AUC with hemisphere (Left/Right), familiarity (Famous/Unknown), and depth (Back/Front) as factors confirmed these results. Both ANOVAs showed a significant main effect of familiarity (M170: $F(1,13) = 5.87, p = 0.031$; M400: $F(1,13) = 6.04, p = 0.029$). Main effects of hemisphere and depth were not significant ($p > 0.2$).

What about the interaction of depth and familiarity? Although famous faces elicited significantly larger responses than unfamiliar faces for both components in the Back condition (M170: $t(13) = 3.63, p = 0.003$; M400: $t(13) = 4.6, p = 0.0005$, paired $t$ tests), but not the Front condition (M170: $t(13) = 0.7, p = 0.5$; M400: $t(13) = 0.57, p = 0.6$), the interaction of familiarity and depth only reached significance for the M400 component (M170: $F(1,13) = 2.02, p = 0.18$; M400: $F(1,13) = 5.29, p = 0.04$). We further verified the results using a non-parametric permutation analysis, which, unlike the parametric $t$ and $F$ tests, is assumption-free regarding the underlying distribution of the data. Again, we found a significant interaction effect for the M400 ($p < 0.04$) but not the M170 ($p < 0.19$).

Another potential concern regarding these results lies in the choice of dependent measure used for each component. While the M170 component is customarily quantified using the amplitude of the peak response, this measure may be less sensitive than the AUC calculation employed for the M400 simply because it relies on fewer...
data points. To ensure that our results do not merely reflect this measurement difference, we also performed an AUC analysis for the M170 response. Despite a significant familiarity effect for the Back ($t(13) = 4.8$, $p = 0.0003$) but not Front ($t(13) = 0.6$, $p = 0.6$), the interaction of depth and familiarity failed to reach significance ($F(1,13) = 1.83$, $p = 0.2$).

This suggests that the difference between the M170 and M400 cannot be explained simply by the greater robustness of the AUC measure. Supporting this point, a follow-up ANOVA on the AUC data with component (M170/M400) as an additional factor showed a significant 3-way interaction of component, familiarity, and depth ($F(1,13) = 6.46$, $p = 0.025$).

Therefore, modulatory effects of familiarity on holistic versus part-based processing occur after the stage of processing indexed by the M170. Supporting this point, Figure 5 displays the depth-by-familiarity interaction across the right MFG from our previous study (Figure 5A), versus the M170 (Figure 5B) and M400 (Figure 5C). Of the two MEG responses, only the later M400 shows a pattern comparable to that seen in the right MFG with fMRI.

Together with our previous fMRI data, our current results from MEG support the idea that both parts and wholes are represented within the face processing stream. If face parts were processed by a more general part-based system, we would expect the responses elicited by the Front depth condition to be smaller than those for faces in the Back depth, and possibly more similar to those for the control category of houses. Instead, we find that both the M170 and M400 show a larger response to faces than to houses across depth conditions, but no significant difference between face depth conditions, suggesting that even part-based representations of faces are coded by the face perception system.

However, there is a possible alternative explanation for our results, especially for the relatively early stage of processing indexed by the M170 response. In comparison to stimuli without binocular disparity cues, our stereoscopic depth manipulations entail additional mid-level visual processing, including resolution of binocular disparity information, assignment of border ownership, and amodal completion. Therefore, it is possible that our failure to find a significantly larger M170 response to the Back (holistic) condition is simply due to the fact that mid-level visual processing is not yet complete. That is, the M170 responses to part-based and holistic conditions could be equivalent for the entirely uninteresting reason that, prior to stereopsis and/or amodal completion, these stimuli are more or less the same.

To address this concern, we directly compared the M170 responses for stereoscopic face stimuli with those for faces without binocular disparity cues, obtained from a separate “localizer” run. This latter data gives us a baseline measure of the M170 response for stimuli without binocular disparity information. Specifically, if the M170 response occurs irrespective of mid-level visual processes such as stereopsis and amodal completion, the latency of this component should be unaffected by the addition of binocular disparity information to the stimulus. In contrast, a significant difference in latency between these conditions would indicate that the M170 occurs after additional mid-level processing.

Figure 5. Comparison of depth-by-familiarity interaction for (A) right middle fusiform gyrus (MFG, circled in yellow at left; from Harris & Aguirre, 2008), (B) M170 amplitude, and (C) M400 AUC. Only the M400 component in MEG shows a similar pattern to that previously measured for the right MFG. The map to the left of each graph depicts the spatial topography of the relevant response, as computed from group average data. Error bars represent standard error of the mean (SEM).
binocular disparity manipulation. Parentheses indicate standard error of the mean (SEM).

<table>
<thead>
<tr>
<th>Condition</th>
<th>Amplitude ($10^{-13}$T)</th>
<th>Latency (ms)</th>
<th>Amplitude ($10^{-13}$T)</th>
<th>Latency (ms)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Face</td>
<td>1.46 (0.12)</td>
<td>209.8 (2.9)</td>
<td>1.01 (0.13)</td>
<td>170.0 (3.0)</td>
</tr>
<tr>
<td>House</td>
<td>0.57 (0.13)</td>
<td>221.2 (5.4)</td>
<td>0.39 (0.16)</td>
<td>155.9 (5.12)</td>
</tr>
</tbody>
</table>

Table 2. Amplitude and latency of the M170 response measured for Face and House stimuli with (StereoFaces) and without (Localizer) binocular disparity manipulation. Parentheses indicate standard error of the mean (SEM).

Table 2 displays the M170 amplitude and latency for faces and houses in the stereoscopic and localizer conditions. Examining the latency of the M170 response for faces, we can see that there is a delay of 41 ms for the stereoscopic, relative to the control, stimuli. In the context of known latency effects, this is a sizable delay: the much-discussed inversion effect in latency is only about 10 ms (Bentin et al., 1996; Itier & Taylor, 2002; Rossion et al., 2000). While effects on the order of 40–50 ms have been reported for isolated nose and mouth stimuli, these are seen in conjunction with reductions in amplitude and broadening of the N170 peak (Bentin et al., 1996; Harris & Nakayama, 2008). The M170 response to faces with our stereoscopic manipulation shows no such decrements, as can be seen in Table 2 and Figure 4. (Indeed, the M170 stereoscopic manipulation shows no such decrements, as can be seen in conjunction with reductions in amplitude and broadening of the N170 peak (Bentin et al., 1996; Harris & Nakayama, 2008). The M170 response to faces with our stereoscopic manipulation shows no such decrements, as can be seen in Table 2 and Figure 4. (Indeed, the M170 stereoscopic manipulation shows no such decrements, as can be seen in conjunction with reductions in amplitude and broadening of the N170 peak (Bentin et al., 1996; Harris & Nakayama, 2008). The M170 response to faces with our stereoscopic manipulation shows no such decrements, as can be seen in Table 2 and Figure 4. (Indeed, the M170 stereoscopic manipulation shows no such decrements, as can be seen in conjunction with reductions in amplitude and broadening of the N170 peak (Bentin et al., 1996; Harris & Nakayama, 2008). The M170 response to faces with our stereoscopic manipulation shows no such decrements, as can be seen in Table 2 and Figure 4. (Indeed, the M170 stereoscopic manipulation shows no such decrements, as can be seen in conjunction with reductions in amplitude and broadening of the N170 peak (Bentin et al., 1996; Harris & Nakayama, 2008).

A repeated-measures ANOVA for M170 latency further revealed a significant main effect of hemisphere ($F(1,13) = 5.0, p = 0.04$), reflecting slightly shorter latencies for the right M170 (208.7 versus 210.8 ms). The main effect of familiarity also approached significance, with longer latencies for familiar relative to unfamiliar faces ($F(1,13) = 4.3, p = 0.06$). Along with the significant main effect of familiarity for amplitude, this is consistent with previously reported familiarity effects at the N170 (Caharel et al., 2006). These data stand in contrast to previous work with famous and unknown faces showing no effect of familiarity at the N170 (Bentin & Deouell, 2000; Eimer, 2000; Schweinberger et al., 2002). The source of this inconsistency is unclear, although there appear to be the case that there is a general latency delay associated with part-based, as opposed to holistic, processing. Instead, these results further support the idea that face parts are important for the relatively early stage of processing indexed by the M170 response.

In addition to supporting our claim that the M170 occurs after mid-level processing is complete, the latency data from our experiment can further inform our interpretation of the results for depth. As mentioned above, one of the main arguments linking the M170/N170 response to configural processing has arisen from the consistent finding of a small (10 ms) but significant latency delay for inverted relative to upright faces, thought to reflect a switch to part-based analysis (e.g., Rossion et al., 2000). If this is indeed the case, we would expect a similar latency delay for the Front (part-based) relative to the Back (holistic) depth condition, given our behavioral finding of “whole-versus-part superiority” (Tanaka & Farah, 1993) for the Back but not Front depth (Harris & Aguirre, 2008).

Figure 6 plots these results along with the non-disparity condition for comparison. In fact, there is no significant difference between the Front and Back depth conditions in terms of M170 latency ($t(13) = 0.16, p = 0.9$, paired $t$ test), although, as described above, responses in both conditions are significantly later than that to the face without binocular disparity information. Therefore, it does not appear to be the case that there is a general latency delay associated with part-based, as opposed to holistic, processing. Instead, these results further support the idea that face parts are important for the relatively early stage of processing indexed by the M170 response.

Figure 6. Mean latency of the M170 response for Front and Back depth conditions and the unoccluded face. If the 10-ms latency delay seen for inversion reflects a switch from configural to part-based processing, we would expect a similar latency delay for the Front condition (right) relative to the Back (middle). Instead, both are later than faces without disparity cues (left), but not significantly different from each other. Error bars represent standard error of the mean (SEM).
are a number of methodological differences between the experiments in question, including the use of ERP versus MEG and the density of the sensor array.

Nonetheless, given that our neurophysiological data show an effect of familiarity, we can further ask whether this effect is correlated with behavioral performance on recognizing faces. Thirteen out of the 14 subjects in the current experiment also completed a familiar/unfamiliar judgment task with a subset of the faces used in the MEG experiment (mean accuracy: 78%).

Of particular interest is the correlation between behavioral performance and MEG response across subjects. We examined the relationship between behavioral recognition performance and the MEG familiarity effect, defined as Familiar - Unfamiliar, for the M170 and M400 in the Back and Front depths. Correlations were computed using the non-parametric Spearman’s rho. The resulting correlations are displayed in Figure 7.

Notably, only the M400 familiarity effect in the Back depth condition is significantly correlated (rho = 0.62, p = 0.023) with behavioral recognition performance. In contrast, even though the M170 shows a significantly larger response to famous versus unknown faces in the Back condition, this familiarity effect is weakly related to behavioral recognition of the face (rho = 0.47, p = 0.1). Likewise, for both the M170 and M400, the amplitude of the response to the part-based Front condition is not associated with recognition performance (M170: rho = 0.2, M400: rho = 0.28, p = 0.35). A permutation analysis confirmed these results, finding a significant correlation between the familiarity effect only for the M400 (p < 0.03) in the Back depth condition (M400 Front: p < 0.37; M170 Back: p < 0.09; M170 Front: p < 0.48). These data therefore lend further credence to the importance of holistic representations in the neural coding of familiar faces.

**Discussion**

Face perception is commonly conceptualized in terms of holistic, as opposed to part-based, processing. Yet, despite the emphasis on face perception as a holistic or configural process (Diamond & Carey, 1986; Farah et al., 1998; Tanaka & Farah, 1993; Young et al., 1987), there is some behavioral and neuropsychological evidence for part-based representations in the face perception stream (Cabeza & Kato, 2000; Leder & Bruce, 1998; Macho & Leder, 1998; Moscovitch et al., 1997).

Recently, we have investigated this question in fMRI using a binocular disparity manipulation derived from Nakayama et al. (1989), in which faces appear either behind or in front of a set of stripes. While the first case allows amodal completion and holistic processing of the face, in the latter the face cannot be completed and is therefore perceived in terms of its parts (Figure 1), as we have demonstrated behaviorally. However, “face-selective” regions of inferotemporal cortex show equal responses to both depth conditions, supporting the idea that both wholes and parts are represented in the face processing system (Harris & Aguirre, 2008).

In keeping with prior behavioral work suggesting that familiar faces may have more robust holistic representations (Buttle & Raymond, 2003; Young et al., 1985), we further found that a region of right MFG previously associated with holistic processing (Schiltz & Rossion, 2006) showed greater adaptation of the response to the holistic versus part-based condition for famous but not unfamiliar faces. Based on these findings, we have argued that “face-selective” regions of cortex are engaged in both holistic and part-based processing, and that the recruitment of part-based versus holistic representations is modulated by familiarity (Harris & Aguirre, 2008).

In the current experiment, we extend these results by characterizing the time course of holistic and part-based processing for “face-selective” responses in MEG. In particular, we examined components in two latency ranges, early (~170–200 ms) and late (~280–480 ms). By testing the responses of these M170 and “M400” components to manipulations of depth and familiarity, we hope to gain a better understanding of when in the face processing stream such effects occur. In particular, based on the prior literature we expected equivalent M170...
responses to the two depth conditions, with familiarity and depth-by-familiarity effects occurring at the later M400 component.

As shown in Table 1, our results largely match these predictions. The M170 shows no significant effects of processing type (holistic versus part-based), as assessed by our depth manipulation. Since this component was measured only at sensors showing “face-selectivity,” the large response in the Front (part-based) condition is unlikely to reflect activity of a separate part-based object recognition system. Rather, these data provide additional evidence for the idea that parts are represented from early in the face processing stream.

Further support for this interpretation comes from the comparison of M170 latency for stimuli with and without binocular disparity cues. A significant 10-ms difference in latency for upright versus inverted faces has previously been argued to reflect a switch from configural to part-based processing. Therefore, we would expect to find similar delays for the Front relative to Back depth, which we have behaviorally linked to part-based versus holistic processing (Harris & Aguirre, 2008). Instead, though processing of stimuli with stereoscopic depth was dramatically slower (~40 ms) than that of unoccluded stimuli (Table 2), there was no significant difference in latency between depth conditions (Figure 6). Consistent with the amplitude data, these results suggest that face stimuli are not differentiated in terms of holistic/configural versus part-based processing at the relatively early stages of face perception indexed by the M170 response. (Note, however, that the present data do not speak directly to the contribution of face parts, as face parts are present in all conditions of this experiment.)

In contrast to the depth manipulation, the effect of familiarity was significant at both the M170 and, as predicted, the later M400 (Figure 4). These data thus replicate recent findings of familiarity effects at the M170/N170 (Caharel et al., 2005; Kloth et al., 2006), unlike previous reports of insensitivity to familiarity (Bentin & Deouell, 2000; Eimer, 2000). The source of this inconsistency is unclear, though it is important to note that these experiments differed on a number of methodological grounds (e.g., task demands, density of sensor array, use of MEG versus ERP).

Interestingly, for both the M170 and the M400, the familiarity effect appeared to be driven by responses in the Back (holistic) condition, as no such pattern was seen for the Front (part-based) condition. Yet only the later M400 component showed a clear interaction of familiarity and depth similar to that seen in fMRI, with a greater response to the Back (holistic) than Front (part-based) depth condition for familiar but not unfamiliar faces (Figure 5). This was confirmed by repeated-measures ANOVA, which found a significant depth-by-familiarity interaction for the M400 response.

Given the lack of depth-by-familiarity interaction for the M170, what causes the significant main effect of familiarity for this response? While we have argued, based on behavioral data, that familiarity increases holistic processing, familiarity could also affect representations in other ways. Indeed, if the M170 indexes a part-based processing stage, as suggested above, the main effect of familiarity could reflect a larger response to familiar, compared to unfamiliar, face parts. Though we cannot exclude the possibility of a small but undetected depth-by-familiarity interaction at the M170, at the very least our data indicate that such interactions, analogous to those seen with fMRI, occur within the first 500 ms after stimulus onset. That the M400, despite its more variable nature relative to the well-defined M170, shows a significant interaction effect moreover suggests that this later component may be particularly sensitive to modulation by familiarity.

Supporting this point, analyses of behavioral data (Figure 7) showed that only the familiarity effect for the Back (holistic) depth in the M400 was significantly correlated with recognition performance. In contrast, familiarity measures in the Back condition at the M170 and in the Front condition at both the M170 and M400 were weakly correlated with behavior. Therefore, it appears that the M400 component, but not the M170, is strongly associated with behavioral recognition, and only for holistically-processed faces. Consistent with existing behavioral data (Buttle & Raymond, 2003; Young et al., 1985), these data further reinforce the importance of holistic representations for familiar faces.

Together with our previous findings from fMRI, the present results more fully elucidate a number of properties of the face processing stream. Most strikingly, in both fMRI and MEG, we have found that faces manipulated in binocular disparity to be perceived either in terms of their parts or wholes nonetheless elicit similar “face-selective” responses. Thus, not only face wholes, but also face parts appear to be represented in the face processing stream. Such part-based representations are unlikely to reflect the activity of a separate part-based object recognition system, as part-based versions of faces evoke significantly larger responses than non-face control conditions in both fMRI (alphanumeric characters; Harris & Aguirre, 2008) and MEG (houses).

The current data extend our work in fMRI by showing that part-based representations are present even at the relatively early stages of face perception indexed by the M170 response. These data are therefore also consistent with our previous work implicating face parts in rapid adaptation of the M170 response (Harris & Nakayama, 2008). Given these prior results, as well as the advantages of using adaptation to assess the M170 response (Harris & Nakayama, 2007), it would be of further interest to probe the M170 response to our “StereoFace” stimuli using this adaptation.

Although “face-selective” responses to part-based face stimuli were seen even for the relatively early M170 component, only the later M400 showed a significant
interaction of familiarity and holistic versus part-based processing, as indexed by depth. While the direct comparison of fMRI and MEG data is not possible due to the differences between these methods in temporal scale, these results suggest a rough temporal window within which modulatory effects of familiarity on processing seen with fMRI may emerge. Though the temporally coarse fMRI response likely integrates other neural interactions such as feedback, the finding of a similar pattern of interaction at the M400 component not only provides an important replication of the fMRI data, but additionally situates such modulatory effects of familiarity within the temporal processing stream.

More generally, these results provide important constraints on accounts of how face and object recognition are instantiated in the human brain. Previous work has often envisioned a rigid dichotomy between face and object recognition, with each system subserved by a separate type of representation. Our data from “face-selective” responses in both fMRI and MEG indicate that this is not the case: for face perception, holistic and part-based representations appear to coexist within a single system. The finding of a depth-by-familiarity interaction further suggests that the relative recruitment of these holistic and part-based representations is not fixed, but rather may be differentially modulated by external factors such as familiarity.

Conclusions

In the present experiment, we used MEG to extend previous findings regarding “face-selective” responses to holistic and part-based representations of faces, and their modulation by familiarity (Harris & Aguirre, 2008). In particular, we examined the responses of two MEG components in early (~170–200 ms) and late (~250–450 ms) latency ranges to stimuli manipulated in binocular disparity to appear behind or in front of a set of stripes (Nakayama et al., 1989), perceived respectively in a holistic or part-based manner. Consistent with our prior data from fMRI, neither the M170 nor the “M400” component recorded at occipitotemporal sensors showed a significant main effect of holistic versus part-based processing, as indexed by depth. In contrast, while both components showed significantly greater responses to familiar faces, only for the later M400 component did this familiarity effect modulate holistic versus part-based processing, with greater responses to the holistic condition for famous but not unfamiliar faces. Collectively, these data suggest that, although part-based representations of faces are present from relatively early in the face processing stream, modulatory effects of familiarity akin to those seen with fMRI occur later in the face processing stream.

Acknowledgments

The authors would like to thank Dr. Timothy Roberts and the MEG Lab of the Children’s Hospital of Philadelphia, and Ranjani Prabhakaram for providing the famous and unfamiliar face stimuli. GKA is supported by a Burroughs-Wellcome Career Development Award and by K08 MH 7 2926-01.

Commercial relationships: none.
Corresponding author: Alison M. Harris.
Email: aharris@alum.mit.edu.
Address: 3 W. Gates Building, Hospital of the University of Pennsylvania, 3400 Spruce St., Philadelphia, PA 19104, USA.

References


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


