Face-Specific Brain Mechanisms: Event-Related Potential (ERP) Evidence Suggesting the Degree of Activation is Modulated by Experimental Features

Rachel Doreen Grills

Abstract: Two major theories regarding the specificity of certain brain regions in the perception of human faces have been put forth. The first, states an existence of a functional neural subsystem dedicated to the processing of human face forms (Farah, 1996). The other argues that this unique brain activity is in actuality a by-product of visual expertise, since similar patterns have been recorded during viewing of other stimuli classes to which the viewer has extensive experience (Gauthier & Curby, 2005). In review of the research, this essay contends that the level of brain activity involved in complex face perception is highly dependent on both image orientation and task demands; however the site of activity origin remains uniform across manipulations. The various experimental methodologies, including basic face-to-object comparisons, inversion effects and feature decomposition according to hierarchical salience, all produce similar results. The prominence of a negative waveform around 170 ms discriminated face processing from both animate and inanimate objects and remained present even with different stimuli orientation. The degree of this activity (i.e., amplitude) and temporal course (i.e., latency to onset) fluctuated as a function of experimental design; however continuity in neural location (i.e., fusiform gyrus and posterior area of the inferior-temporal) was maintained.

The highly evolved human brain has many specialized regions dedicated to processing certain types of information (Kolb & Whishaw, 2001). With regards to the perception of human faces, a unique area of the brain dedicated to processing physiognomic information has been proposed (De Hann, Humphreys & Johnson, 2001; Farah, 1996). Researchers that endorse this phenomenon state an innate existence of a domain-specific functional subsystem localized to the right hemisphere of the posterior temporal and inferior occipitotemporal cortex (Bentin, Allison, Puce, Perez & McCarthy, 1996; Farah, 1996). The opposing theory, states that the neural activity thought to solely reflect cognitive processing of human faces, is in actuality, a function of visual expertise. This perspective asserts that the aforementioned brain regions can be recruited for the detection of homogeneous object classes if the observer is highly proficient in the particular field of interest (Gauthier & Curby, 2005; Rossion, Kung & Tarr, 2004).

To expand upon the latter presumption, that face processing is simply part of a global processing mechanism, event-related potential (ERP) evidence has revealed an acute decrease in face-specific activity after the comparison group of homogenous objects were trained to a level of expertise (Rossion et al., 2004). This means that when comparing the neural activity evoked during face viewing, to that elicited during the viewing of trained objects, the difference no longer reached a level of distinction. Likewise, Gauthier and Curby (2005) exemplified a similar effect of expertise when testing car enthusiasts, during the viewing of various automobiles. They showed that these car experts had greater activity levels in the brain regions thought to be reserved for facial processing (Gauthier & Curby, 2005). Taken together, these two lines of evidence suggest that the localized brain activity generated during face processing is in actuality a by-product of visual expertise (Gauthier & Curby, 2005; Rossion et al., 2004). Such findings imply that these brain areas can be recruited for cognitive processes aside from face perception and moreover that stimuli proficiency governs patterns of neural activation.

Despite these findings, overwhelming evidence is in favor of particular brain areas dedicated to processing both holistic face representations as well as adjusted configurations. Support for these neural regions has been derived from a variety of sources. For one, activity recorded from single-cell electrodes implanted in the brain of monkeys indicates that face-specific cells maximally fire only when the animal responds to human faces (Desimone et al., 1984). This preferential response is recorded in the inferotemporal cortex and superior temporal sulcus (Bruce, Desimone & Gross, 1981). These face-cells fire around 100-200 milliseconds (ms) when differentiating the face form, but are also capable of specifically responding to the physical characteristics, such as the eyes, as well (Perrett, Mistlin & Chitty, 1987). Secondly, studies with prosopagnosia patients indicate deficits in recognizing familiar human faces but not objects. This has been evidenced in patients with either right unilateral or bilateral lesions of the inferior occipitotemporal regions and parietotemporal cortex (Buchtel, 2001; Damasio, Tranel & Damasio, 1990; Farah, 1996). Along similar lines, Buchtel (2001) showed that for unilateral right temporal lesions a face recognition deficit was present regardless of which visual field the image was presented to. This evidence builds the foundation for a neural processing mechanism regulated to respond specifically to human faces.

With technological advancements, many of the current evaluations of this face-specific neural phenomenon have revolved around electrophysiological measures such as ERP recordings as well as imaging techniques, explicitly functional magnetic resonance imaging (fMRI) and positron emission topography (PET). PET research suggests that facial processing elicits augmented cerebral blood flow in right fusiform gyrus and right inferotemporal cortex (Rossion, Schiltz, Robaye, Pierenne & Crommelinck, 2001). Likewise, fMRI research has revealed augmented activity levels in the right fusiform gyrus, reaffirming
The Comparison of Faces to Non-Face Stimuli

An elementary experimental paradigm used to examine facial processing simply compares a collection of images comprised of faces and other classes of objects, while simultaneously recording brain activity during viewing. A frequently utilized electrophysiological measure in the assessment of brain activity patterns are ERPs. Due to their excellent temporal resolution, they provide valuable information on the timing of particular events (Kolb & Whishaw, 2001). The N170, a negative potential elicited between approximately 150-200 ms after stimulus onset, is thought to reflect specialized face processing activity (Eimer, 2000; Sagiv & Bentin, 2001). Electrophysiological recordings have indicated the presence of a focially distributed N170 over the lateral posterior scalp during the viewing of faces but not objects (Bentin, Allison, Puce, & McCarthy, 1996; Eimer, 2000) This negative, high amplitude waveform was reserved to the viewing of human faces and not elicited to any other classes of animate and inanimate stimuli (Bentin et al., 1996; Eimer, 2000; Herrmann, Ehls, Muehberger & Fallgatter, 2005; Itier & Taylor, 2004). The N170 remained insensitive to other body parts, such as human hands, suggesting that this negativity captures neural processing attuned to human faces and not simply global visual identification (Bentin et al., 1996; McCarthy et al., 1999). By expanding the signature face ERP's to voxel by voxel comparisons of source localized neural sites, researchers have been able to pinpoint particular regions in the brain thought to generate this unique activity. In particular, the maximal difference between faces and objects reveals heightened N170 activity in the fusiform gyrus and extending to parieto-temporo-occipital cortex (Herrmann et al., 2005; Itier & Taylor, 2006). Segmentation mapping, done with the computerized analysis of electrophysiological data, has indicated a separate cluster of brain activity during the viewing of faces (Itier and Taylor, 2006). This reinforces the unique neural quality of facial processing when compared to complex object analysis.

Additional ERP evidence for face-specific processing is derived from the P100, a positive potential at around 100 ms, which appears during comparisons of faces to objects (Herrmann et al., 2005; Itier & Taylor, 2004; McCarthy et al., 1999). This potential shows greater amplitude and shorter latency when comparing facial images to those of objects (Eimer, 2000; Itier & Taylor, 2004) and modified faces (Halit et al., 2000). With regard to localization, magnetoencephalography (MEG) research showed activity residing over the occipitotemporal cortex (Liu, Harris, & Kanwisher, 2002). The activity was slightly posterior to the characteristic potential evoked at 170 ms (Liu et al., 2002); however still indicating that face processing occurs more quickly than object processing. The P100 was thus interpreted to be a representative marker of the early and more elementary stages of facial processing (Herrmann et al., 2005; Liu et al., 2002).

Another ERP noted in previous research is the N200. This negative potential appears approximately 30 ms later than the N170, in the inferior temporal gyri, lateral to the occipitotemporal sulcus during the viewing of faces (Herrmann et al., 2005). Given the similarities in location and preferential response to faces, the N170 and N200 are proposed to represent similar cognitive processes (Bentin et al., 1996; Herrmann et al., 2005). The differences in timing may be a by-product of methodological differences and across subject averaging. In summary, the N170 is the most prominent marker for specialized facial processing and is thought to reflect structural encoding processes (Eimer, 2000; Rossion et al., 1999) while the N200 and the P100 are supplementary potentials that reaffirm the evidence in favour of neural facial processing mechanisms. For simply holistic, face-to-object comparisons, these signature ERPs are thought to reflect the underlying neural mechanisms dedicated to processing physiognomic information (Herrmann et al., 2005). The most prominent N170 marker, notably localized to the fusiform gyrus and parieto-temporal-occipital region, indicates that face processing recruits a specialized, interrelated neural network specific to physiognomic characteristics.

Inverted faces Relative to Upright Faces

Another way to assess if the brain is hard-wired to process faces in a unique manner is through comparing the same stimuli presented in different orientations. One approach used, is the evaluation of regional brain activity patterns when the traditional facial alignment is distorted. The common distortion is facial inversion where images are rotated to an upside-down position. The brain activity in response to these inversions is then compared to that elicited during upright viewing. A collection of objects which serve as a control group are then used to gauge the similarities and differences between the two face positions.

Relative to a control montage comprised of animate and inanimate objects, both inverted and upright facial alignments elicited the N170 (Bentin et al., 1996; Eimer, 2000; McCarthy et al., 1999; Rossion et al., 1999). The distinction between the inverted and upright conditions was the emergence of a notable delay in the inverted placement, although activity was still constituted within the range of the N170 (Bentin et al., 1996; Hann et al., 2002; Eimer, 2000; Itier & Taylor, 2002; Sagiv & Bentin, 2001). Furthermore, inversion has been found to impact N170 amplitude characteristics. Relative to the upright faces, inverted images revealed augmented N170 amplitudes; however this effect was not evidenced for inverted control objects (Bentin et al., 1996; Eimer, 2000; Itier & Taylor, 2004; Sagiv & Bentin, 2001). Alterations to the N200 have also been revealed in response to inverted face stimuli. These changes mirror the amplitude and latency attributes of the N170 to virtually the same degree (McCarthy et al., 1999).

The P100 is also altered by manipulating inversion techniques. As previously mentioned, the P100 is representative of more elementary face processing stages. It is less discriminatory activity form elicited for both the upright and upside-down conditions; however a more pronounced positive waveform with an extended
duration is shown with inversion (Bentin et al., 1996; Itier & Taylor, 2004). Specifically, the external differences between the two face alignments elicited longer latencies with greater amplitudes for the inverse figures (Eimer, 2000; Herrmann et al., 2005; Itier & Taylor, 2004). The significance of this evidence lies in the underlying neuronal control mechanisms for properties of face identification. It indicates that global facial characteristics are detected by neurons very early (100ms) and that this premature response may be holistic in nature enabling attention mechanisms to concentrate on these important biological features.

Activity was consistent for the same region despite the external manipulations of face position (Itier & Taylor, 2004). The bilateral negativity was uncovered in the parietal-occipital cortex regions and remained invariant for all faces, but was not present during the object presentations (Herrmann et al., 2005). An interesting phenomenon for inversion was discovered by McCarthy et al. (1999) when the inverted images were examined in terms of contralateral and ipsilateral visual field presentations. When images were presented to contralateral visual fields, the right hemisphere (RH) produced a more rapid and strong N200 for the upright face whereas the left hemisphere (LH) showed the same for the inverted face. In opposition, when the stimuli appeared in the ipsilateral visual field, the pattern was reversed. The RH now revealed an elevated N200 amplitude and corresponding shorter latency in response to the inverted face, whereas the LH showed an equivalent pattern for the upright face (Herrmann et al., 2005).

Taken together, the similar localization of activity for both facial positions is congruent with the assumption that a neural mechanism attuned to the structural features of faces exists around the posterior inferotemporal region. However, the notable alteration in the degree of activity reaffirms the argument that the specific methodological characteristics of facial images can shape temporal processing characteristics. Lastly, it was identified that the hemisphere laterality plays a role in inversion perception. Specifically the RH reveals greater proficiency for processing information about upright faces and transferring that material to the LH whereas the left hemisphere specializes in reconstructing inverted faces (Herrmann et al., 2005; McCarthy et al., 1999). Suggestively, the RH advantage often attributed to face processing may be due to a holistic spatial perception, since the image configuration is in fact. When this configuration is distorted by inversion, the LH may be recruited for piece by piece reconstruction.

### Facial Features (Eyes) Relative to Entire Face Compositions

A recent theoretical assumption regarding neural specificity for faces states that localized brain activity may not be a result of the face per say, but instead a by-product of activity for particular components of the face especially the eyes (Bentin et al., 1996; Farah, 1996; Jemel et al., 1999; McCarthy et al., 1999; Sagiv and Bentin, 2001). Given the similarity of the N170 elicited for both inverted and upright faces, researchers have attempted to separate and evaluate components of feature salience. Behavioural research has previously indicated that eyes are the most pronounced feature (. Face parts are ranked in level of salience from eyes to mouth and then nose, from highest to lowest respectively (Shepherd et al., 1981; as cited in McCarthy et al., 1999). Subsequently ERP research has expanded upon this evidence in terms of electrical sensitivity to the prominence of the various facial features. Indeed, findings reveal that this feature salience is interrelated to the degree of the N170 produced. The N170 was elicited only for whole faces and eyes, with greater waveform amplitudes allocated to presentations of the eyes in isolation (Bentin et al., 1996; Eimer, 2000; Jemel et al., 1999). Likewise, McCarthy et al. (1999) showed that ERP activation patterns followed an identical order from greatest to least facial feature salience for the N200 potential. All face parts were shown to evoke an N200 relative to non-face objects; however, the magnitude of this activation was modified by the salience of the eyes. Heightened amplitudes were shown in response to eyes shown in isolation, slightly exceeding those engendered in conventional face presentations (McCarthy et al., 1999). This suggests that the neural generator producing the N170 may not be reliant on face integrity, but instead on the presence and clarity of the eyes (Bentin et al., 1996). To test the importance of eyes, Bentin et al. (1996) altered the context in which the eyes appeared. Even in fragmented face contexts, a significant prominence of the N170 prevailed. Therefore, it is conceivable that face processing is not purely holistic in nature and reasonable to conclude that the presence of the N170 marker may be more indicative of salient feature identification of the eyes, rather than specific to face recognition as a whole (Bentin et al., 1996; McCarthy et al., 1999).

The specific topography of brain activity in response to the eyes has also been investigated. Using an associative priming paradigm Jemel et al. (1999) examined the ERPs produced during face processing with the identification of either congruent or incongruent eyes superimposed on a familiar face. The participants were asked to identify if the eyes shown on the face were consistent or different than their expectations derived from the eyes-blanked prime. The results revealed a spatiotemporal overlapping for face-to-feature congruency. The primed face elicited a vertex P200 as well as an N170 in the inferotemporal region. Moreover, the latencies were reduced for both the N170 and P100. Such shortened latencies indicated that in the absence of the eyes, the face was encoded quicker but not to its full extent. In contrast, an N210 peaked bilaterally with the presentation of eyes superimposed in the face. Moreover, this peak was unique to the unexpected condition, in which the participants subjectively perceived the eyes as incongruent to their anticipation formed by the prime. This activity increase was localized to posterior inferotemporal region. Thus, I suggest that the brain mechanism attuned to processing human faces may be influenced by cognitive expectations potentially indicative of a top-down influence on interpretation. When expectations or schemes are violated, top-down organization is disrupted. Hence, a greater level of brain activity is recruited to the face region to re-process the visual information using bottom-up reconstruction to compensate for the cognitive violation.

Research by Eimer (2000) found that the hierarchical salience of the eyes was not the only feature that affected ERPs during face processing. The N170 was delayed and attenuated for faces lacking internal features such as the nose and mouth and external features such as the face outline. This suggests that this negative face-processing marker may have a wide range of sensitivity and its scope may be morphed in accordance to numerous facial components (Eimer, 2000). It could also indicate that the N170 is representative of the latter stage of the structural encoding process, since global face perceptions need to be constructed for subsequent recognition and memory (Butchel, 2001; Eimer, 2000; Farah, 1996).
Less Predominantly Examined Variables in Face Processing Research

Other variables of interest have been tested with regards to their role in characteristic face processing activity. One less frequently examined factor is the influence of colour. Grayscale versus colour images were compared by McCarthy et al. (1999) and neither amplitude nor latency varied as a function of the colour parameters. Both face representations equally evoked the N200 suggesting that this ERP is generated in response to the face configuration as opposed to the pigment of the facial figure (McCarthy et al., 1999). With regards to size of the face image, both enlarged and shrunken pictures elicited the same ERP effect irregardless (McCarthy et al., 1999). In contrast, directionality of face position (full-face, three quarter and profile views) changed the amplitude of the N200 with the profile view eliciting a significantly smaller potential to the turned-head conditions (McCarthy et al., 1999). Likewise, Eimer (2000) showed a distinct attenuation of the N170 potential during the viewing of cheek and back profiles relative to the full-face alignment. With regards to scrambling and blurring of facial images, greater N170 amplitude (Bentin et al., 1996) and increased N200 latency (McCarthy et al., 1999) have been observed for the jumbled and hazed conditions respectively. Lastly, variations in face attractiveness and setting of the face, either in natural or artificial contexts, have also been manipulated (Halit et al., 2000). This experimental modulation reduced the P100 amplitude in the artificial setting, without changing the latter N170 waveform. With respect to the natural condition both the P100 and N170 potentials were larger in amplitude and evidenced a delayed latency. This indicates that early face perception is more easily modified by exogenous changes in the expected form, whereas latter markers of face perception depend more significantly on the exact nature of the stimuli and implicit task demands (Halit et al., 2000).

In conclusion, ERP research indicates that there are specialized mechanisms (represented by electrical activity generated at particular temporal timelines) attuned to both faces and face components. When faces are compared to objects, electrical activity is elicited in the occipito-temporo-parietal region especially in the right hemisphere and fusiform gyrus, with the signature marker being the N170 event. As mentioned earlier, it was assumed that the actual characteristics of the faces may influence the degree of electrical activity elicited. The effects of face manipulation were deconstructed in terms of inversion, feature salience with a focus on the role of the eyes and basic comparisons to both animate and inanimate control objects. Faces were discriminated from objects in terms of the N170 evoked from scalp electrodes around the posterior region of the inferior-temporal cortex and fusiform gyrus. Additional distinctions in face processing functions were shown with inversion and presentation of the eyes. Both delayed the neural response mechanisms, and augmented the overall amplitude of brain activity especially with respect to the negative potentials elicited around 170 ms. The crucial point for interpretation, is that despite these changes in the levels of activity and temporal pattern, the site of origin in the brain remained astoundingly consistent across the various manipulations. Therefore, it can be concluded that although effect size fluctuates with the experimental design and componental features of the face itself, a neural region specialized for facial processing still appears to be responsible for the unique activity associated with face viewing. This does not preclude or eliminate the influence of visual expertise in the neural activity elicited to face perception.

Faces in general are one of the most frequently viewed and interpreted visual images in day-to-day existence (Gauthier & Curby, 2005). Thus, there could be a secondary or mediating influence of expertise given the dynamic nature of facial processing. It simply suggests that unique, localized activity may be a product of both the face itself and potentially past experiences. An interesting research paradigm would test participants’ memory for holistic and fragmented faces (emphasizing in particular the eyes) across sleep conditions using electrophysiological recordings. By simply adding emotion to some of the facial stimuli, one could measure the effects of emotion on declarative memory consolidation during rapid eye movement sleep to determine if face reprocessing occurs during that sleep stage (or any others). At a later interval, memory could be measured using a behavioural test, and ERPs could be recorded a second time to delineate any changes in brain activity patterns due to prior learning and memory. Such research could further elucidate higher-order cognitive processes of face perception, while separating and clarifying the role of past learning and experience.

References


Itier, R.J., & Taylor, M.J. (2004). N170 or N1? Spatiotemporal differences between object and face processing using ERPs. *Cerebral Cortex, 14,* 132-142.


