Face-sensitive cortical processing in early infancy

Hanife Halit, Gergely Csibra, Ágnes Volein, and Mark H. Johnson
Centre for Brain and Cognitive Development, School of Psychology, Birkbeck College, University of London, UK

Background: Debates about the developmental origins of adult face processing could be directly addressed if a clear infant neural marker could be identified. Previous research with infants remains open to criticism regarding the control stimuli employed. Methods: We recorded ERPs from adults and 3-month-old infants while they watched faces and matched visual noise stimuli. Results: We observed similar amplitude enhancement for faces in the infant N290 and adult N170. In contrast, the infant P400 showed only a latency effect, making it unlikely to be the main precursor of the adult N170. Conclusions: We conclude that there is some degree of specificity of cortical processing of faces as early as 3 months of age. Keywords: Event-related potentials, N170, infants, face perception.

Evidence from single-cell recordings in monkeys, human adult neuropsychology, and functional neuroimaging on adults converge to suggest that visual images of faces are processed by specialised neural and cognitive mechanisms that differ from those involved in the processing of other objects (Farah, 1996; Tovee, 1998, but see Gauthier, Tarr, Anderson, Skudlarski, & Gore, 1999). At present it is unclear how such specialised processing of faces arises during development. While some have suggested that there is an innate domain-specific neural circuit for face processing (Farah, Rabinowitz, Quinn, & Liu, 2000), others have argued that face specialisation results from perceptual expertise (Gauthier, & Nelson, 2001). This issue has important implications for our understanding of developmental disorders in which aspects of face processing are affected (e.g., Dawson et al., 2002; Grice et al., 2001, in press).

While evidence from behavioural studies has been useful in testing hypotheses about the typical and atypical development of face processing, issues related to neural specialisation can most directly be addressed through functional neuroimaging. For a variety of reasons functional neuroimaging with healthy awake infants is difficult. A recent positron emission tomography (PET) study with 2-month-old infants showed that pictures of unfamiliar faces activated a distributed network of cortical areas, including areas activated in adults (the fusiform gyrus), as well other areas not activated in adult studies (Tzourio-Mayoyer, De Schonen, Crivello, Reutter, & Aujard, 2002). However, the control stimulus used to obtain the subtraction images (a moving dot array) differed from the face stimuli in many ways, including being dynamic (rather than static), and differing in spatial frequency and contrast. Several other laboratories have used scalp-recorded event-related potentials (ERPs) to determine the specificity or otherwise of cortical processing of faces over the first year (see de Haan, Johnson, & Halit, 2003 for review). In these studies two types of stimuli have been used to compare with faces: (1) pictures of objects such as familiar and unfamiliar toys (de Haan & Nelson, 1999), and (2) face-related stimuli such as inverted faces and monkey faces (see de Haan et al., 2003 for review). Although at first sight objects such as toys may seem like ideal comparison stimuli to faces, such stimuli are generally not matched for low-level psychophysical properties such as spatial frequency. This is a particular concern for studies with infants where such factors are known to influence preferences and processing (Banks & Salapatek, 1981). In addition, potentially graspable objects may engage activity in other neural circuits concerned with action planning (Kaufman, Mareschal, & Johnson, 2003). While the use of inverted faces overcomes many of these problems, uncertainty remains about the regularity with which young infants view faces in the vertical axis. To date, no study with infants (or arguably, with adults) has compared faces to ideally matched visual stimuli. We note that parallel issues persist in the infant speech perception literature (Vouloumanos & Werker, 2004). In the present experiment we compare faces to ‘visual noise’ stimuli created to match both spatial frequency and colour spectra while recording ERPs from infants and adults.

Event-related potentials recorded from adults have consistently shown that faces elicit a negative potential that occurs most prominently over occipito-temporal scalp regions and peaks between 140 and 170 ms after stimulus onset (Bentin, Allison, Puce, Perez, & McCarthy, 1996). This ERP component, termed the N170, is sensitive to faces in that it is of larger amplitude and shorter latency to faces compared with most other stimuli, including hands, feet, trees, cars, letters and words (Botzel & Grusser, 1989; Botzel, Grusser, Haussler, & Naumann, 1989; Bentin et al., 1996; George, Evans, Fiori, Davidoff, &
Renault, 1996). Moreover, the amplitude and/or latency of the N170 also differs between upright human faces and closely related stimuli such as monkey or ape faces and inverted human faces (Eimer, 2000c; Rebai, Poiroux, Bernard, & Lalonde, 2001; Carmel & Bentin, 2002; de Haan, Pascalis, & Johnson, 2002). Whilst the functional significance of the N170 in adult face processing has yet to be fully elucidated, it is thought to reflect processing related to the structural encoding of faces as opposed to recognition of facial identity (Eimer, 2000a, b). This conclusion is based on findings demonstrating that the N170 is sensitive to experimental manipulations that alter the basic structural properties of faces, e.g., inversion, feature scrambling (George et al., 1996; Eimer, 2000a, b), but is unaffected by the familiarity (Bentin & Deouell, 2000; Eimer, 2000a) or emotional expression of faces (Eimer & Holmes, 2002).

Since the N170 is considered to be an electrophysiological marker of specialised mechanisms for face processing, knowledge about its development during postnatal growth can be key to our understanding of the development of cortical specialisation for face processing. Recent attempts to identify the developmental precursors of the N170 have involved examining the morphology, spatial distribution, timing and functional specificity of ERP components elicited by faces in infants over the first year of life (de Haan & Nelson, 1999; de Haan et al., 2002; Halit, de Haan, & Johnson, 2003). These studies identified two components over the posterior scalp region that were sensitive to manipulation of faces: a negative-going deflection (N290) that peaked at 290 ms after stimulus onset, and a positive-deflection (P400) that peaked at 400 ms after stimulus onset.

The primary aim of the present study was to investigate the cortical processing of faces in young infants with an ideally matched control stimulus. A secondary aim was to assess which of two candidate ERP components (N290 or P400) is most likely to form the precursor to the adult N170. To determine this, we recorded ERP responses to faces and to matched visual noise control stimuli in adults (Experiment 1) and 3-month-old infants (Experiment 2).

**Experiment 1**

**Materials and methods**

**Participants.** Twelve volunteers (7 female, mean age 28.6 years) took part in the study. All participants were right handed and had normal or corrected to normal vision.

**Stimuli.** The stimuli consisted of 25 colour images of human female faces and 25 matched visual noise images that served as a control. Noise stimuli were created by randomising the phase spectra of the faces whilst keeping the amplitude and colour spectra constant1 (see Figure 1). All stimuli were presented against a white background and subtended a horizontal angle of 15° and a vertical angle of 21° when viewed from a distance of 60 cm. Participants were presented with 150 trials of each stimulus in random order. Stimulus duration was 413 msec and the inter-stimulus interval varied randomly between 800 and 1200 msec.

**Procedure.** Participants were asked to fixate the centre of the computer screen where the stimuli appeared but were not required to perform any particular task. The stimuli were presented in a single block.

**EEG Recording.** Continuous EEG was recorded using a Geodesic Sensor Net (GSN) consisting of 128 silver-silver chloride electrodes evenly distributed across the scalp and referenced against the vertex electrode (Tucker, 1993). The GSN also includes electrodes next to, and below, the eyes for recording horizontal and vertical eye-movements. The signal was amplified by an EGI NetAmps amplifier (Eugene, OR), with a bandpass filter of .1–100Hz. The continuous EEG was digitised at 250 Hz sampling rate and was stored on a computer disk for off-line analysis.

**ERP Waveform Analysis.** The continuous EEG was divided to create segments from 200 ms pre-stimulus onset to 800 ms post-stimulus onset. Data were manually edited for electrical and movement artefacts and then digitally filtered offline with a 30Hz low-pass elliptical filter. The entire trial was excluded if data from more than 10 sensors were excluded or if the trial contained an eye-blink. Data were baseline-corrected against the average voltage during the 100 ms prior to stimulus onset and average ERPs were created

---

1 Specifically, (1) a two-dimensional Fast Fourier Transformation was applied to all three colour components of the images, (2) the phase on each frequency was replaced by a random value between $-\pi$ and $\pi$ (uniform distribution), (3) an inverse Fourier transformation reconstituted the image, (4) to which the outer contour of the original face was applied as a mask.
Results and discussion

Figure 2C depicts grand average ERPs at the posterior temporal sites. The N170 was of larger amplitude for faces \( (M = -2.14 \mu V, SD = 2.2 \mu V) \) compared with noise \( (M = -21 \mu V, SD = 1.87 \mu V) \) (main effect of Condition: \( F(1,11) = 66.30, p < .001 \)). In contrast, the peak latency of the N170 did not differ between faces and noise. Further effects showed that the N170 was of larger amplitude (main effect of Site: \( F(1,11) = 5.29, p < .05 \)) and peaked earlier (main effect of Site: \( F(1,11) = 19.64, p < .01 \)) over lateral (Mamp = -1.54 \( \mu V, SD = 1.97 \mu V; \) Mlat = 174.75 ms, SD = 11.63 ms) compared with medial electrode sites (Mamp = -0.81 \( \mu V, SD = 2.19 \mu V; \) Mlat =163.83 ms, SD = 14.21 ms). We did not find a Hemispheric effect (see Figure 2B).

The lack of hemispheric differences in the amplitude of the N170 is consistent with previous studies of face processing (Eimer, 1998; Itier & Taylor, 2002; Rossion et al., 2000).

The results obtained in Experiment 1 confirmed that the N170 was larger amplitude for faces compared with the noise stimuli. This result is consistent with previous studies that have compared faces with other non-face stimuli and is also consistent with the notion that the adult N170 is an electrophysiological marker for specialised mechanisms for the processing of faces.

Experiment 2

Materials and methods

Participants. The final sample consisted of thirteen 3-month-old infants (8 males) with an average age of 85.4 days (range 77–93 days). We selected 3-month-olds since this is the youngest age at which most theorists agree that face processing is mediated by the cortical circuits that give rise to ERPs (de Haan et al., 2003). All infants were born full term and were of normal birthweight. An additional 34 infants were tested but were excluded from further analysis due to eye and/or body movements that resulted in recording artefacts \( (n = 30) \) or due to refusal to wear the net \( (n = 4) \). This exclusion rate is comparable to other studies using electrophysiological measurements in young infants (e.g., de Haan et al., 2002).

Stimuli. We used the same stimuli as in Experiment 1.

Procedure. After the 64-channel sensor net was applied, each infant passively viewed the stimuli while seated on the carer’s lap in a dimly lit booth, approximately 60 cm from a 21 inch-computer monitor mounted in a black background. Both infant and carer were readily observable to the experimenter at all times via a video camera situated beneath the monitor. Stimuli were presented only when infants were seen to be watching a central fixation point and were withdrawn when the infant became too fussy or bored to attend. Between trials the experimenter could activate a noise via a speaker also located beneath the presentation monitor and/or presented coloured patterns on the monitor to re-orient the infant’s attention to the screen if required. The average number of trials viewed for faces and noise was 70 (range: 34–134).

ERP waveform analysis. The general method of ERP analysis was the same as in Experiment 1. Additionally, video recordings of each infant’s behaviour during the test session were coded offline to ensure that they were fixated centrally to the stimuli at the time of presentation and that trials containing artefacts caused by vertical or horizontal eye-movements or eye-blinks could be eliminated. The mean number of trials included in the average ERPs was 31 for faces and 28 for visual noise. A paired samples t-test found this difference to be statistically significant \( (t(12) = 3.2, p < .01) \). This difference occurred as a consequence of infants’ increased separately for faces and noise for each participant. Missing data were interpolated using spherical spline interpolation from the individual subject averages. ERPs were re-referenced to the average reference.

Statistical analyses of the data were carried out on the peak amplitude and peak latency of the N170 at four different recording sites covering occipito-temporal scalp regions: right medial (around O2), left medial (around O1), right lateral (around T6), and left lateral (around T5), each site including 6 electrodes (see Figure 2A). The amplitude and latency of the N170 were derived by establishing the peak amplitude and peak latency within a specified time window (120–200 ms) for the individual electrodes, and averaging them within electrode sites. These measures were analysed in a 2×2×2 repeated measures ANOVA with Condition (face, noise), Hemisphere (left, right) and Site (lateral, medial) as within-subject factors. An alpha of .05 was used in all analyses.
tendency to look away from the noise stimuli compared with faces, thus causing a greater number of trials to be rejected due to eye-movement artefacts.

Grand-averaged data were visually inspected and two time windows were chosen to capture the face-sensitive components identified in previous infant studies: (a) N290 (180–450 ms) and (b) P400 (450–550 ms). All individual averages were also visually inspected to ensure the chosen time windows were appropriate. Similarly to the analysis of the adult data, the peak amplitudes and latencies were aggregated into four regions representing the left and right, and lateral and medial, occipito-temporal cortices (see Figure 3A). Because the infant EEG was recorded by 64-channel nets, only 3 electrodes fell into each of these regions. These measures were analysed the same way as in Experiment 1.

Results and discussion

Both face and noise stimuli elicited a negative component that peaked maximally approximately 330 ms after stimulus onset (N290), and was followed by a positive component that peaked around 500 ms (P400) after stimulus onset (see Figure 3C). Means and standard deviations (SD) of peak amplitudes and latencies of the N290 and P400 are shown in Table 1.

**N290.** Concerning the N290 amplitude, all three main effects were significant. The amplitude of the N290 was larger for faces compared with noise ($F(1,12) = 30.52, p < .001$), was larger over the right than over the left hemisphere ($F(1,12) = 10.21, p < .01$), and was larger medially then laterally ($F(1,12) = 4.80, p < .05$). No interactions were found significant for the amplitude of the N290. The analysis on the peak latency of the N290 did not yield any significant effect.

**P400.** The amplitude of the P400 did not differ between faces and noise. In contrast, the latency of the P400 peaked earlier for noise compared to faces (main effect of Condition, $F(1,12) = 10.84, p < .01$). No further main effects or interactions were significant.

The results of Experiment 2 showed that the amplitude of the P400 was not sensitive to the difference between face and noise stimuli in 3-month-olds, while the amplitude of the N290 displayed a huge difference. In contrast, the peak latency of the P400 discriminated between the stimuli, but, unlike in the adult studies that demonstrated latency effects on the N170, it was earlier to non-face than to face stimuli. These results suggest that the functional characteristics of the N290 are more similar to the adult N170. Note also that the effect demonstrated in this experiment was much more robust than the ones reported in earlier studies that compared ERPs to upright and inverted faces. The robustness of the present effect can be seen both in its duration (the difference between the ERPs to faces and noise started as early as the P1 peak and lasted until the P400) and in the wide distribution over the posterior cortex.

General discussion

The primary aim of this study was to investigate the cortical processing of faces in young infants with an ideally matched control stimulus. The present result, when taken together with previous ERP results comparing faces to objects, establishes conclusively that by 3 months there is some degree of specialisation of cortical processing of faces. While this
relatively early specialisation for face processing initially appears consistent with the innate domain-specific module view, a number of other lines of evidence show that further development takes place during the subsequent months and years. First, cortical processing of faces at 3 and 6 months appears to be less specialised (narrowly tuned) than in 12-month-olds and adults (de Haan et al., 2002; Halit et al., 2003). Second, in both infants (Tzourio-Mazoyer et al., 2002) and children (Passarotti et al., 2003) face processing activates a greater extent of cortical tissue than in adults. Both of these lines of evidence are consistent with an 'interactive specialisation' perspective on human functional brain development in which the functionality of cortical regions arises as a result of interactions between brain regions, and between the whole brain and its environment (Johnson, 2000; Johnson, Halit, Grice, & Karmiloff-Smith, 2002). From this perspective, some developmental disorders may be described as delayed or failed specialisation (Johnson et al., 2002, Grice et al., in press).

A secondary aim of this study was to assess which of two candidate ERP components (N290 or P400) is most likely to form the precursor to the adult N170. There are now several studies that have investigated ERP components related to face processing during infancy and childhood. For example, Taylor, McCarthy, Saliba, and Degiovanni (1999) investigated the N170 in children between 4 and 14 years of age and found that whilst the N170 was identifiable at all ages it was of smaller amplitude and longer latency compared with the adult N170. Further studies addressed the question of the development of face-sensitive components during the first year of life. In one study, de Haan et al. (2002) investigated the development of cortical specialisation for face processing by comparing 6-month-old infants’ and adults’ ERP responses while they passively viewed upright and inverted human and monkey faces. They found that the adult N170 showed specificity to upright human faces in that: (1) it was of smaller amplitude and shorter latency for upright human faces compared with all other face conditions and (2) face inversion increased its amplitude and latency for human faces only (de Haan et al., 2002). An equivalent ERP component to the adult N170 was not identified in infants. Rather, sensitivity to species and orientation of faces appeared to occur at two separate components within 6-month-old infants’ ERP waveform. Firstly, the N290, a negative-going deflection occurring over posterior scalp regions, peaked at 290 ms after stimulus onset. Like the adult N170, the N290 showed sensitivity to the species of faces in that it differed in amplitude between monkey and human faces. In contrast to the adult N170, the N290 did not show sensitivity to the orientation of faces. The N290 was followed by a positive-deflection, the P400, also occurring most prominently over posterior electrodes that peaked at 400 ms after stimulus onset (see also de Haan & Nelson, 1999). Like the adult N170, the P400 was influenced by inversion: its amplitude was more negative for inverted than upright faces. However, unlike the N170, this effect was not specific to human faces and also occurred for monkey faces. Thus, these findings showed that neither the N290 nor the P400 exhibited the same degree as specificity for upright human faces as the adult N170.

In a follow-up study, Halit et al. (2003) investigated whether either of the two face-sensitive infant components (N290 and P400) could be considered as developmental precursors of the adult N170. They used the same stimuli as de Haan et al. (2002) to investigate the effects of species and orientation in 3- and 12-month-old infants. At 12 months of age the N290, like the N170, showed sensitivity to both the species and orientation of faces: (1) it differed between human and monkey faces and (2) inversion increased the amplitude of the N290 for human faces but not monkey faces. The P400 also showed specificity to human faces in that inversion delayed its latency for human faces but not monkey faces. Both an increased amplitude and delayed latency are a characteristic response of the adult N170 to human face inversion. This was not found to occur at 6 months of age (de Haan et al., 2002) and 3-months (Halit et al., 2003). Since the response properties of both the N290 and P400 showed and increased specificity to human faces with age, neither could have been ruled out as a precursor to the adult N170.

The present study attempted to resolve this question by directly comparing ERPs to faces with ERPs to non-faces while keeping certain low-level psycho-physical properties equal. We measured the electrophysiological responses to these stimuli at the youngest age when the measurement of transient evoked responses is feasible. Our results suggest that, at a functional level, the N290, as opposed to the P400, is more likely to correspond to the developmental precursor of the N170.

Previous studies have shown that the P400 is also face-sensitive in infants (de Haan & Nelson, 1999; de Haan et al., 2002; Halit et al., 2003). The finding that the P400 did not differ in amplitude between faces and noise in our study does not necessarily rule out the possibility that it reflects processing related to face processing. In adults, the N170 is thought to reflect structural encoding of faces (Eimer, 2000a, b), whilst other aspects of face processing, e.g., emotion analysis, familiarity judgements, are thought to be reflected at components that occur beyond the N170 (Eimer 2000a). One possibility is that the absence of task requirements, necessary when testing young infants, did not engage processing mechanisms that would be typically engaged in making emotional or identity judgements about faces and hence P400 processing was minimal in this study. There are currently no
studies that have investigated facial identity recognition in infants and adults within a single study. If, like adults, infants show sensitivity to the various aspects of face processing at different ERP components then this suggests that infants may be using similar mechanisms to process faces as adults.

The question of whether the neural generators of the infant N290 are the same as the N170 remains to be determined. The N170 is thought to reflect the activation of one or more of a network of temporal lobe face processing regions, including regions within the fusiform gyrus (Shibata et al., 2002), the posterior temporal gyrus (Bentin et al., 1996; Shibata et al., 2002), superior temporal sulcus, and lateral occipito-temporal cortex (Bentin et al., 1996; Schweinberger, Pickering, Jentzsh, Burton, & Kaufmann, 2002). The larger N170 to faces versus noise observed in the present study may reflect activation of one or more of these areas. Interestingly, the amplitude of the N290 was larger to faces than to noise, suggesting that cortical areas underlying the processing of faces are, at least in part, active at 3 months of age.

When taken together with the results of other recent studies, the infant N290 shows a discriminatory response to: (1) faces versus matched visual noise, (2) human versus monkey faces (Halit et al., 2003), and (3) faces with direct versus averted gaze (Farroni, Csibra, Simion, & Johnson, 2002). Despite these similarities to the adult N170, the infant N290 also undergoes significant changes in its response properties during postnatal development. In terms of cortical specialisation, the type of changes seen to occur at the N290 are consistent with increased functional specialization (or ‘perceptual narrowing’) of its underlying neural generators (de Haan et al., 2002).

Acknowledgement

This work was supported by the UK Medical Research Council Programme Grant No. G9715587.

Correspondence to

Gergely Csibra, Birkbeck College, Malet Street, London WC1E 7HX, UK; Tel: +44 20 7631 6323; Fax: +44 20 7631 6587; Email: g.csibra@bbk.ac.uk

References


Manuscript accepted 1 March 2004