One versus two eyes makes a difference!

Early face perception is modulated by featural fixation and feature context

Karisa B. Parkington [a]* and Roxane J. Itier [a]

[a] Department of Psychology, University of Waterloo, Waterloo, ON, Canada

*corresponding author

Department of Psychology

University of Waterloo

200 University Avenue West

Waterloo, ON, Canada

N2L 6H9

Karisa B. Parkington's email address: kparkington@uwaterloo.ca

Roxane J. Itier's email address: ritier@uwaterloo.ca

ABSTRACT

The N170 event-related potential component is an early marker of face perception that is particularly sensitive to isolated eye regions and to eye fixations within a face. Here, this eye sensitivity was tested further by measuring the N170 to isolated facial features and to the same features fixated within a face, using a gaze-contingent procedure. The neural response to single isolated eyes and eye regions (two eyes) was also compared. Pixel intensity and contrast were controlled at the global (image) and local (featural) levels. Consistent with previous findings, larger N170 amplitudes were elicited when the left or right eye was fixated within a face, compared to the mouth or nose, demonstrating that the N170 eye sensitivity reflects higher-order perceptual processes and not merely low-level perceptual effects. The N170 was also largest and most delayed for isolated features, compared to equivalent fixations within a face. Specifically, mouth fixation yielded the largest amplitude difference, and nose fixation yielded the largest latency difference between these two contexts, suggesting the N170 may reflect a complex interplay between holistic and featural processes. Critically, eye regions elicited consistently larger and shorter N170 responses compared to single eyes, with enhanced responses for contralateral eye content, irrespective of eye or nasion fixation. These results confirm the importance of the eyes in early face perception, and provide novel evidence of an increased sensitivity to the presence of two symmetric eyes compared to only one eye, consistent with a neural eye *region* detector rather than an eye detector *per se*.

Keywords: N170, eye sensitivity, eye region, face perception, neural inhibition

Highlights

- N170 modulations to isolated features and features fixated in a face were examined.

- The N170 eye sensitivity reflects high-order perceptual processes.

- Isolated features yield larger and slower N170s than features fixated in a face.

- Two eyes elicit larger N170s, of intermediate latency, versus single eyes and faces.

- A neural eye region detector mechanism during early face perception is proposed.

1. INTRODUCTION

The eyes convey a wealth of cues that facilitate a variety of social behaviours (e.g., social interaction, inter-personal connection, mutual gaze and social attention; Itier, 2015; Itier & Batty, 2009; Kleinke, 1986). A wide body of behavioural and eye-tracking studies have shown that, while idiosyncratic differences in face scanning patterns are evident both across and within participants (Kanan, Bseiso, Ray, Hsiao, & Cottrell, 2015; Mehoudar, Arizpe, Baker, & Yovel, 2014; Miellet, Caldara, & Schyns, 2011), neurotypical adults spend more time, on average, attending to the eye region of a face than to the nose and mouth (Birmingham, Bischof, & Kingstone, 2008a, 2008b; Haig, 1985; Henderson, Williams, & Falk; 2005; Janik, Wellens, Goldberg, & Dell'Osso, 1978; Yarbus, 1967; for reviews see Itier, 2015 and Shepherd, Davies, & Ellis, 1981). This attention bias towards the eye region is thought to maximize the perception and extraction of basic social cues. Attention to, and use of, eye information has indeed been shown to improve face identity, gender, and emotional expression judgements (Haig, 1985; Hills, Cooper, & Pake, 2013; Schyns, Bonnar, & Gosselin, 2002; Vinette, Gosselin, & Schyns, 2004). Thus, the eyes contain unique information whose successful extraction seems necessary for accurate higher-order social cognitive processing.

The importance of the eyes has also been demonstrated at the neural level using scalp-recorded event related potentials (ERPs). In particular, the N170 ERP component, a neural marker of early face perception, is recorded maximally over occipito-temporal regions between 120-200 ms post-stimulus onset (Bentin, Allison, Puce, Perez, & McCarthy, 1996; George, Evans, Fiori, Davidoff, & Renault, 1996; Rossion & Jacques, 2011). The N170 is generally considered a neural marker of holistic face perception (Eimer, 2000a, 2000b; Rossion & Gauthier, 2002; for a review see Rossion & Jacques, 2011), a stage during which features are perceptually "glued" into an

indecomposable whole. However, a growing body of literature indicates that the N170 is also sensitive to eye information.

In particular, isolated eye regions elicit increased N170 amplitudes and delayed latencies compared to faces (Bentin et al., 1996; Itier, Alain, Sedore, & McIntosh, 2007; Itier et al., 2006; Itier, Van Roon, & Alain, 2011; Kloth et al., 2013; Taylor, Itier, Allison, & Edmonds, 2001; Cauquil, Edmonds, & Taylor, 2000; Shibata et al., 2002). This N170 eye sensitivity, which seems to be more pronounced for human eyes compared to eyes of other species (Itier et al., 2011; Shibata et al., 2002), is also present within the context of a full face, such that fixation on the eyes or nasion yield larger N170 amplitudes relative to fixation on other features (de Lissa et al., 2014; Itier & Preston, 2018; McPartland, Cheung, Perszyk, & Mayes, 2010; Neath & Itier, 2015; Neath-Tavares & Itier, 2016; Nemrodov, Anderson, Preston, & Itier, 2014). Moreover, the N170 is also sensitive to eye colour and eye shape (Zheng et al., 2011), and the detection of contralateral eye information seems to be coded during the transition between the P100 and N170 peaks (Rousselet, Ince, van Rijsbergen, and Schyns, 2014). Collectively these findings support the view that the human brain is attuned to the early detection and processing of the eyes within a face, and suggest the presence of a neural eye detector mechanism.

Critically, the existence of an eye sensitivity within a face opposes the generally accepted view that faces are processed in a purely holistic manner, and instead suggests that early face perception involves a combination of featural and holistic processes. In an attempt to integrate these opposing views, Nemrodov et al. (2014) proposed the Lateral Inhibition, Face Template, and Eye Detector (LIFTED) model of early face perception. In essence, the LIFTED model proposes that early face perception, as marked by the N170 ERP component, involves a complex

interplay between foveal[1] and parafoveal information, both of which influence the holistic percept (see also Itier, 2015). The model suggests that a neural eye detector mechanism first enables the detection of the eyes which serve as anchor points, from which the position and distances to the other features are coded based on an upright human face template. If the typical upright face configuration is detected (i.e., two eyes above a nose, above a mouth, within an oval outline), a neural inhibition mechanism kicks in, wherein the neurons coding for parafoveal information inhibit the neurons coding for the fixated feature, to allow for equal representation of all face parts. If the normal face configuration is not detected (e.g., if the face is inverted or if isolated features are shown), the inhibition mechanism is not triggered, and the N170 would then be larger than when recorded to a whole (upright) face, as it would represent the combined activation related to the processing of the foveal and parafoveal contents. In this framework, the N170 represents the combined activity related to the processing of the eye feature and to a holistic processing stage mediated by a neural inhibition mechanism.

The goal of the current study was to clarify the impact of these featural and holistic processes and further examine the role of the eyes in early face perception. First, we sought to re-examine the neural responses to isolated facial features. Specifically, the N170 response to facial features other than the eyes is largely understudied, with most investigations focusing primarily on the isolated eye region. In addition, studies that investigated the neural response to facial features typically compared features to whole faces centred on the nasion or nose. However, given the evidence that featural fixation within a face modulates N170 responses (de Lissa et al., 2014; Itier & Preston, 2018; McPartland et al., 2010; Neath & Itier, 2015; Neath-Tavares & Itier, 2016; Nemrodov et al., 2014), the comparison of isolated features to faces should be made when fixation

---

[1] Here, fovea is defined as extending ~1º of visual angle on either side of fixation (for a total of 2º of visual angle), parafovea as extending 1-5º eccentricity, and periphery as extending beyond 5º (Polyak, 1941; Rodieck, 1998).

location is controlled. Therefore, the current study compared the N170 response to the same facial features presented in isolation and within the context of a face, using a gaze-contingent procedure. By keeping a given feature in fovea, any N170 modulations recorded between the isolated feature and the same feature within a face context can be attributed to the presence of parafoveal information in the face context. As previously reported with eye regions (e.g., Bentin et al., 1996), we predicted that one single eye, a condition not tested before, would still yield a larger N170 amplitude and faster latency relative to isolated mouths or noses, reflecting a special sensitivity to the eye compared to other facial features. Based on the LIFTED model, we also predicted larger and later N170 responses for isolated features, compared to features fixated within a face, due to the neural inhibition mechanism triggered when viewing upright whole faces. This inhibition, or face context effect, was expected to be larger for nose and mouth fixations compared to eye fixations because, in accordance with the LIFTED model, the eyes should only be partially inhibited (which is why fixation on the eyes of a face elicit the largest N170 responses in the first place; see Nemrodov et al., 2014 for an in-depth discussion).

We also sought to address the question of whether the N170 eye sensitivity within a face is driven by low-level visual properties or rather reflects a higher-order perceptual process. Human eyes are a zone of high contrast (Kobayashi & Kohshima, 1997), and although global (image) low-level factors such as mean contrast and pixel intensity are known to modulate the N170 component (Rossion & Jacques, 2011), their impact at the local (featural) level is, to the best of our knowledge, unknown. The current study addressed this issue by equating the internal facial features (left eye, right eye, nose, and mouth) on pixel intensity and contrast. We predicted that the eye sensitivity would still be seen using these locally controlled stimuli, thereby demonstrating that this sensitivity

is not merely due to differences in these low-level image properties, but instead reflects a higher-level perceptual process.

The final goal of the present study was to refine our understanding of the neural response to the eyes. In particular, we addressed the question of whether the eyes might be special because there are two symmetrical eye features (compared to only one nose feature or one mouth feature). To this end, eye regions (two eyes) were compared to single eyes while gaze position was enforced on an eye or on the nasion (i.e., the region between the two eyes). The presentation of isolated facial features (eye region, one eye, one nose, or one mouth) disrupts holistic processing; however, the classic view, according to which any disruption of the holistic processing triggers a larger N170 amplitude compared to whole faces (e.g., Rossion & Jacques, 2011), does not make any prediction as to whether a difference should be found between the eye region and one eye. In fact, it could be argued that, with two symmetrical eyes, the eye region is perceptually closer to a face than one single isolated eye, with a somewhat intermediate level of holistic disruption. Within this classic framework, the N170 recorded to the eye region should be of an intermediate amplitude and latency compared to that recorded to a face and that recorded to one single eye, and should thus be smaller and earlier for the eye region than for one single eye. In contrast, the LIFTED model predicts a larger N170 to the eye region than to a single eye. Indeed, according to the model, the neural inhibition mechanism should not be triggered for either one isolated eye or for the eye region, as the face configuration is disrupted in both cases. The neural activity recorded should thus represent the combined neural activity elicited by foveal and parafoveal information, i.e. the combined activity elicited by each of the two eyes for eye regions. This cumulative (rather than inhibitory) process should be seen irrespective of whether an eye or the nasion is fixated.

2. MATERIALS AND METHOD

2.1 Participants

Thirty-four undergraduate students participated in the current study (19 females, 25 right-eye dominant, M = 21.29 years, SD = 3.04). All participants reported normal or corrected-to-normal vision, had lived in Canada and/or the United States for at least ten years (M = 19.32 years, SD = 3.33), did not have a history of head injury, neurological disease, psychiatric disorder, epilepsy, seizures, or sensitivities to flashing lights, and were not taking medications containing cortisone (which could increase sensitivity to infra-red light emitted by the eye-tracker) nor anti-psychotic medications. Two additional participants did not complete the experiment due to difficulties tracking their eyes, and one participant registered too few trials due to a large number of eye movements and other artifacts. This study was approved by a University of Waterloo Human Research Ethics Committee. In accordance with the Declaration of Helsinki, all participants provided informed written consent at the beginning of the experiment, and participated for Psychology course credit.

2.2 Stimuli

Thirty grey-scale identities (15 male, 15 female) were selected from the face database produced by Nemrodov et al. (2014), initially generated using FACES™ 4.0 by IQBiometrix Inc. All internal facial features (left eye, right eye, nose, and mouth) were equated on mean pixel intensity[2] and root mean squared (RMS) contrast[3] at the local level. Four circular regions of interest (ROIs) were created (1.92º diameter) centered on each feature (Figure 1B), and the pixel intensity and RMS contrast values for each of these ROIs were adjusted using custom MATLAB

---

[2] Calculated as the mean value for all pixels within the featural ROI (local) or across the entire image (global).
[3] Calculated as the standard deviation value for all pixels within the featural ROI (local) or across the entire image (global).

scripts, so that the mean values were virtually identical across features (Table 1). This pixel adjustment used smoothing algorithms to avoid the creation of darker areas around the features. All stimuli were mirror-flipped along the vertical axis to eliminate asymmetry effects, thereby doubling the number of images.

Table 1. Mean pixel intensity and RMS contrast across the pre-defined 1.92º regions of interest (ROIs) centered on each feature. Standard deviations are in parentheses.

Featural ROI	Pixel Intensity	RMS Contrast
Left Eye	0.62 (0.002)	0.09 (0.001)
Right Eye	0.62 (0.002)	0.09 (0.001)
Nose	0.62 (0.004)	0.09 (0.002)
Mouth	0.62 (0.005)	0.10 (0.003)

Each identity was subsequently cropped into an oval (8.13º x 12.64º), so as to only include the internal features of the face (Figure 1). Isolated feature stimuli were created by extracting, from each facial oval, a rectangular segment (6.92º x 2.45º) centered on the eye region, the nose, or the mouth. Single isolated left/right eye stimuli were generated by airbrushing out one eye from the eye region stimuli. All face and feature stimuli were then placed on a pixel-scrambled background (12.88º x 17.89º; Figure 1) in such a way that the eye fixations were centred on each pupil, and fixations on the nasion (the area between the two eyes), the tip of the nose and the mouth were aligned vertically along an axis passing through the midline of the face. All stimuli were modified using the GNU Image Manipulation Program (GIMP 2.8), and were further adjusted at the global (image) level to be equivalent in mean pixel intensity (0.58) and RMS contrast (0.48). This was accomplished using homemade MATLAB scripts based off the SHINE toolbox (Willenbockel et al., 2010), which equalized the mean pixel intensity and RMS contrast of the whole image by only changing properties of the pixelated background. This was done to ensure pictures were equated

at the global level, without affecting the local properties of the actual face part of the image, which were already equated locally within the desired ROI. Using this two-step process, the stimuli were equalized in both mean pixel intensity and mean RMS contrast at the global image level and at the local face ROI level, without affecting the quality of the face itself (no extra light/dark zones or loss of image quality).

Fourteen experimental conditions (5 face conditions and 9 isolated feature conditions) were created based on stimulus type (isolated feature or face) and fixation location (left eye[4], right eye, nasion, nose, or mouth; Figure 1A): i) full face-left eye fixation (FF-LE), ii) full face-right eye fixation (FF-RE), iii) full face-nasion fixation (FF-Na), iv) full face-nose fixation (FF-No), v) full face-mouth fixation (FF-M), vi) isolated eye region-left eye fixation (IEyes-LE), vii) isolated eye region-right eye fixation (IEyes-RE), viii) isolated eye region-nasion fixation (IEyes-Na), ix) isolated nose-nose fixation (INo), x) isolated mouth-mouth fixation (IM), xi) isolated left eye-left eye fixation (ILE-LE), xii) isolated left eye-nasion fixation (ILE-Na), xiii) isolated right eye-right eye fixation (IRE-RE), and xiv) isolated right eye-nasion fixation (IRE-Na). Furthermore, oddball and practice stimuli were created by inverting all nine isolated feature conditions and three of the face conditions (left eye, right eye, and nasion) for one extra face identity.

---

[4] Please note that eye position will always be referenced from the observer's perspective (i.e., *left eye* is the eye on the left side of the image, and *right eye* is the eye on the right side of the image).
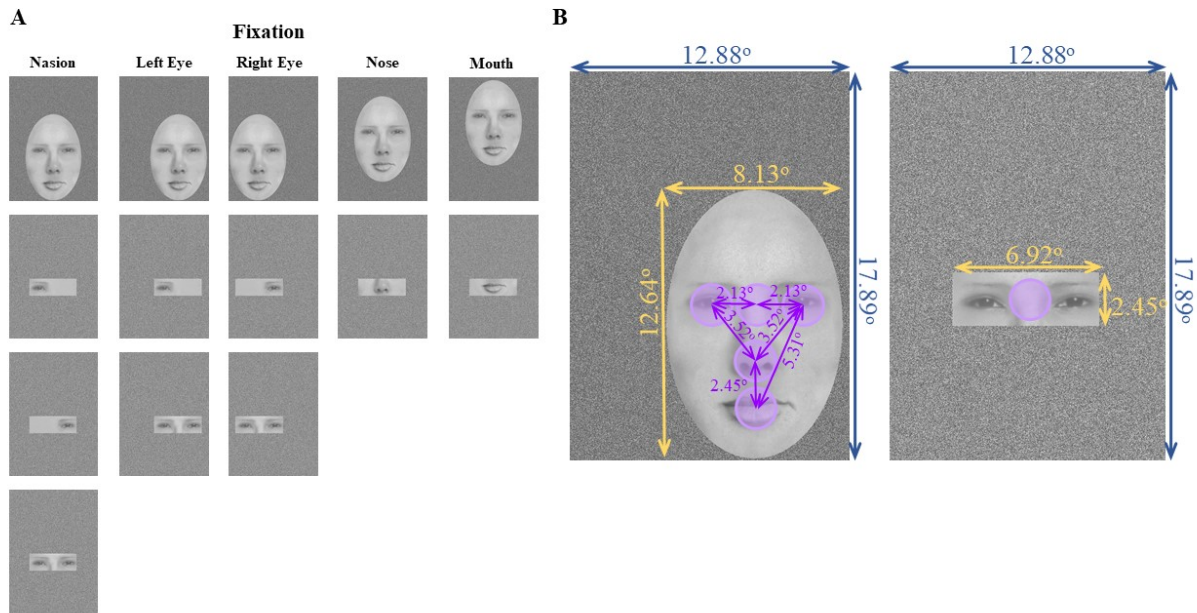
Figure 1. *Panel A:* Examples of the 14 face and isolated feature conditions. Participants always fixated on the same location on the computer screen, and each stimulus was presented offset to ensure the participant's gaze landed on one of the five fixation locations (nasion, left eye, right eye, nose, or mouth). Note that eye positions are referenced from the observer's perspective (i.e., left eye indicates the eye on the leftmost side of the image). *Panel B:* Stimulus exemplars (face with left eye fixation, and an eye region with nasion fixation) with image size and angular distances between fixation locations. Purple circles represent the non-overlapping 1.92º ROIs centered on each fixation location, which were used to reject eye movements, and were equated for mean pixel intensity and RMS contrast.

2.3 Design

The study was conducted in a dimly-lit sound-attenuated Faraday cage, with participants seated 70 cm away from the computer screen, and their heads supported in a chinrest. Participants were told to focus on a fixation cross, and that their gaze would trigger the appearance of a picture of a face or facial feature. Participants were further instructed to maintain this fixation when the picture appeared (i.e., no eye movements). Throughout the experiment, participants completed an oddball detection task (10% probability)[5], by pressing the spacebar only to inverted faces and

---

[5] Each experimental block consisted of fourteen oddball trials: nine isolated feature trials (IEyes-LE, IEyes-RE, IEyes-Na, INo, IMo, ILE-LE, ILE-Na, IRE-RE, and IRE-RE), two FF-LE trials, two FF-RE trials, and one FF-Na trial.

features. All participants completed 18 practice trials before the experimental blocks. Five participants repeated the practice phase due to eye-tracking difficulties or to clarify instructions.

An Eyelink 1000 desk-mounted remote system (SR Research, http://sr-research.com), sampling at 1000 Hz, was used to measure eye position. At the beginning of each block the participant's dominant eye (as determined by the Miles test; Miles, 1930) was calibrated using a nine-point automatic calibration sequence. The non-dominant eye was recorded for three participants due to calibration issues with the dominant eye. Each experimental trial began with a gaze-contingent fixation trigger, during which participants had to maintain fixation for 300 ms within an ROI subtending 0.98º x 0.98º visual angle centred on a cross (presented 15.95º horizontally and 11.86º vertically from the top-left corner of the computer monitor). Due to variability in maintaining gaze position constant across trials, the average duration of this fixation trigger across participants was 553 ms (SD = 128 ms). This trigger activated stimulus presentation for 250 ms, followed by a response screen that remained until a button press was made, or for a maximum of 700 ms (Figure 2), whichever criterion was met first. If the fixation trigger was not activated within ten seconds (i.e., if the participant failed to fixate on the fixation cross for 300 ms), the trial was aborted and a drift correct was recorded. Mid-block re-calibrations occurred following three sequential drift corrects or when the eye recording was clearly off-centre.

Each block of trials contained 140 experimental stimuli (10 stimuli – 5 of each gender – for each of the 14 conditions) and 14 oddball trials. Twelve blocks were included in this experiment, resulting in a total of 1848 trials overall (120 trials per condition x 14 conditions = 1680 experimental trials + 168 oddball trials). However, due to time limitations, participants completed an average of 9 blocks (SD = 2). The experiment was programmed under Experiment

Builder 1.10.1385 (SR Research, http://sr-research.com) and stimuli were presented on a 1600 x

1200 pixel CRT computer monitor with a refresh rate of 85 Hz.
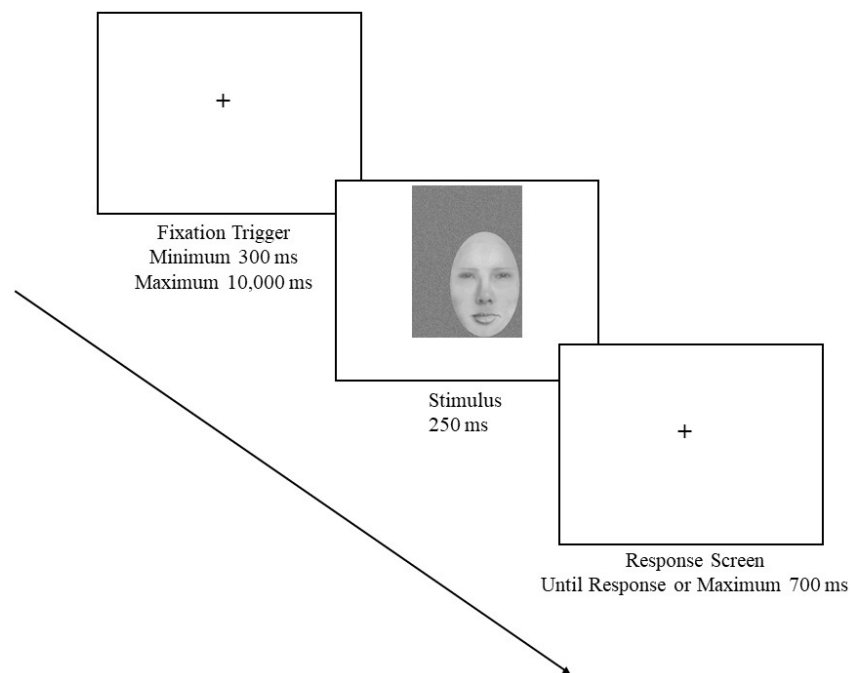

Figure 2. Trial example with left eye fixation on a face. First, the fixation cross was displayed on the screen as a fixation trigger for a minimum of 300 ms and a maximum of ten seconds ($M = 553$ ms $\pm$ 128). If the trigger was not activated within this time period, the trial was aborted and a drift correct was initiated. Once the fixation trigger was activated, the stimulus was presented for 250 ms, followed by a response screen (with a fixation cross) for a maximum of 700 ms. Responses were only required for oddball trials (i.e., infrequent inverted stimuli).

2.4 Electroencephalogram (EEG) Recordings

EEG recordings were acquired using an Active-two Biosemi system with custom-made 72-electrode caps: 66 channels in an electrode cap under the 10/20 system extended (including PO9 and PO10 electrodes over the occipito-parietal region), and three pairs of face electrodes (two pairs of electrodes situated on the outer canthi and infra-orbital ridges to monitor horizontal and vertical eye movements, and one pair of electrodes situated over the mastoids). A Common Mode Sense (CMS) active-electrode and a Driven Right Leg (DRL) passive-electrode served as a ground during

13

recording[6]. EEG data were collected continuously at 512 Hz, and electrode offset was kept below a value of ±20 mV.

2.5 Data Processing

Only upright face/feature trials for which no response was made were included in the present analyses (0.28% of experimental trials were removed across participants due to misses and false alarm responses). EEG and ERP data were analyzed using EEGLab (Delorme & Makeig, 2004) and ERPLab (http://erpinfo.org/erplab) toolboxes in MATLAB.

All EEG data were average-referenced offline, and eye-tracking and EEG recordings were synchronized to ensure the maintenance of continuous fixation during the fixation trigger and stimulus presentation. Trials in which fixations extended beyond the 1.92° ROI for the centered feature were excluded from further analysis, resulting in removal of 3.38% of trials across participants. Data were digitally band-pass filtered (0.01 Hz – 30 Hz) and epoched into time segments extending from -100ms pre- to 350ms post-stimulus presentation, using the pre-stimulus time as a baseline. Trials with artifacts above or below ±70µV were automatically detected and removed, and additional manual cleaning was conducted for one participant due to remaining artifacts, resulting in exclusion of 12.92% of trials across all participants. After trial rejection, participants provided an average of 76 trials per condition ($SD = 23$).

---

[6] Please note that the Biosemi ActiveTwo system does not have an online reference during recording acquisition through the use of active electrodes (https://www.biosemi.com/faq/cms&drl.htm).
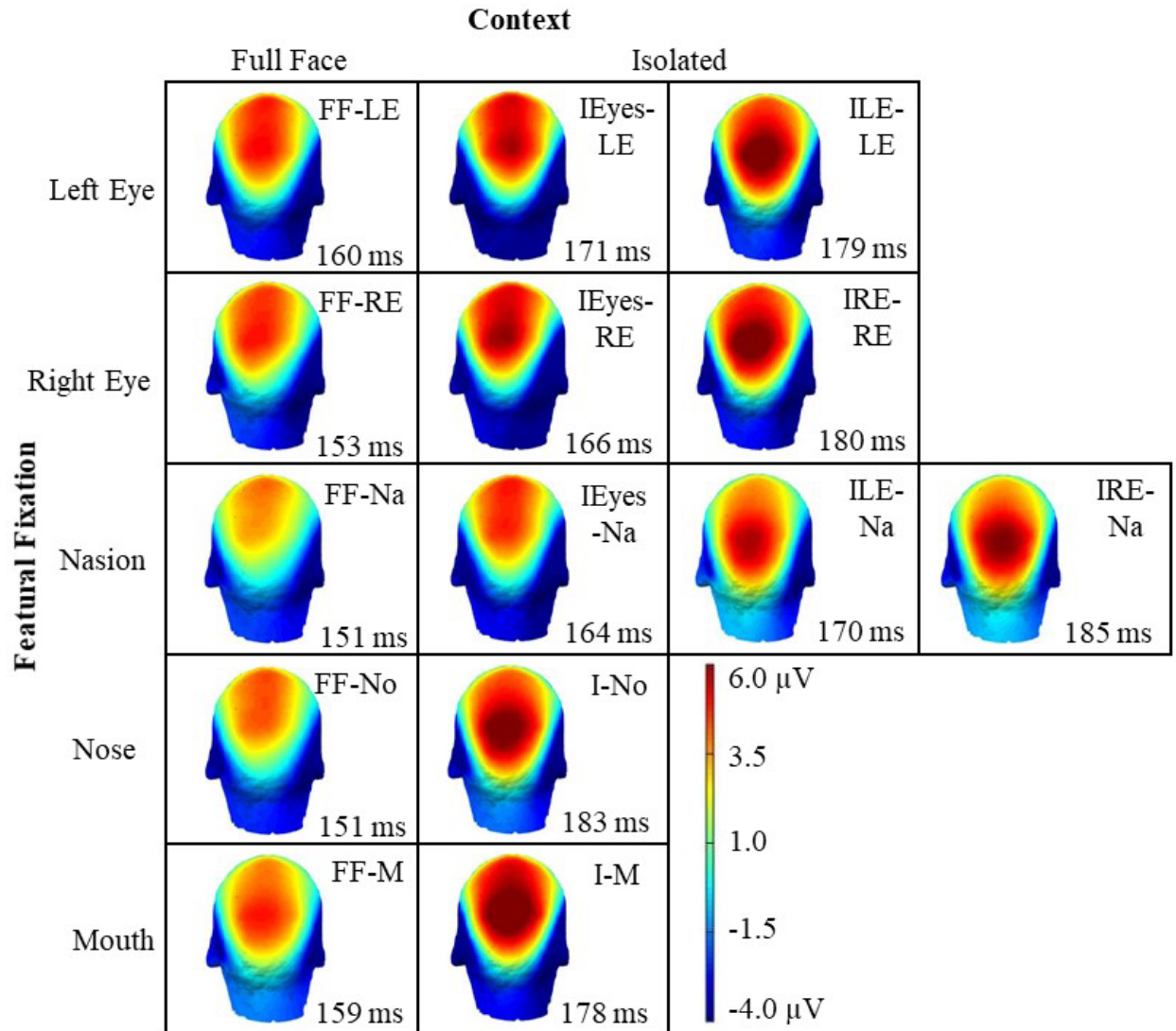
2.5 Data Analysis

Figure 3. Scalp topographic maps generated for each experimental condition, at the latency of the N170 peak recorded for the group averaged ERP on the right hemisphere (back-of-the-head views).

As shown in Figure 3, all conditions demonstrated classic N170 topographies during face and feature viewing. This distribution was observed for all participants with, notably, one electrode within each hemisphere demonstrating strongest peak N170 responses for all conditions. These two electrodes (one per hemisphere) were determined individually for each participant. This

approach was taken, rather than the commonly used grand average approach, in order to be maximally sensitive to individual differences in N170 peaks (see also Itier & Preston, 2018; Neath & Itier, 2015; Neath-Tavares & Itier, 2016; Itier & Neath-Tavares, 2017). The distribution of selected peak electrodes is presented in Table 2. The peak N170 amplitude and latency were measured at these electrodes between 120 ms and 215 ms post-stimulus onset.

Table 2. Number of participants who showed a maximal N170 peak at left and right hemisphere electrode locations.

Left Hemisphere		Right Hemisphere	
Electrode	# Participants	Electrode	# Participants
		P8	2
P7	1	PO8	2
PO9	6	PO10	11
P9	27	P10	19

All analyses employed repeated measures analyses of variance (ANOVAs) conducted on the N170 peak amplitudes and latencies for a subset of conditions, depending on the question. For clarity, we describe the ANOVA used at the beginning of each analysis result section. All analyses were carried out using SPSS Statistics 23. Greenhouse-Geisser adjusted degrees of freedom were applied when Mauchley's Test of Sphericity was violated ($p < .05$), and Bonferroni corrections were applied to all paired comparisons.

3. RESULTS

3.1. Isolated Features vs. Features Fixated Within a Face Context

Modulations of the peak N170 amplitude and latency to isolated features and features fixated within a face were analyzed separately using repeated-measure ANOVAs with Featural Fixation (4: left eye, right eye, nose, mouth), Context (2: full face, isolation), and Hemisphere (2:

left, right) as within-subject factors. The following conditions were included: FF-LE, FF-RE, FF-No, FF-M, IEyes-LE, IEyes-RE, INo, and IM (Figure 1A). This analysis tested the following predictions: i) the N170 should be larger for isolated features compared to the same features fixated within a face context; ii) this context effect should be smaller for eyes than nose or mouth; iii) the N170 eye sensitivity reported for faces should be replicated despite the internal facial features being equalized for pixel intensity and RMS contrast; and iv) an eye sensitivity should also be seen for isolated features with a larger and faster N170 for single isolated eyes.

3.1.1 N170 peak amplitude.

The main effect of Context was significant, such that fixation on isolated features elicited consistently larger N170 amplitudes than fixation on the same features within a face context ($F(1,33) = 26.92$, $MSE = 28.24$, $p < .001$, $\eta_p^2 = .45$; Figure 4A). Amplitudes were also largest over the right hemisphere (main effect of Hemisphere: $F(1,33) = 9.48$, $MSE = 87.18$, $p = .004$, $\eta_p^2 = .22$). This right hemispheric lateralization was most pronounced for left eye, right eye, and nose fixations, and least pronounced for mouth fixation, as revealed by a Hemisphere by Featural Fixation interaction ($F(3,99) = 11.85$, $MSE = 2.26$, $p < .001$, $\eta_p^2 = .26$; Figure 4A-C). The main effect of Featural Fixation ($F(2.16, 71.31) = 9.14$, $MSE = 3.10$, $p < .001$, $\eta_p^2 = .22$) was also significant, but was further qualified by a Context by Featural Fixation interaction ($F(3,99) = 27.18$, $MSE = 2.65$, $p < .001$, $\eta_p^2 = .45$; Figure 4D-E).

N170 context difference scores (isolated feature amplitude – face amplitude) were calculated for each featural fixation. Analysis revealed that mouth fixation yielded the largest context difference of 4.53 µV (ps $< .001$), followed by eyes and nose fixations which did not differ from each other (ps $= 1.00$; Figure 4E). When the face context was analyzed separately (Figure 4B, 4D), left and right eye fixations yielded larger amplitudes compared to mouth (ps $< .001$) and

nose (left eye: $p = .001$; right eye: $p = .04$) fixations, and the amplitude for nose fixation was in turn larger than for mouth fixation ($p = .001$; simple effect of Featural Fixation: $F(3,99) = 25.90$, $MSE = 2.51$, $p < .001$, $\eta_p^2 = .44$)[7]. Alternatively, when isolated features were analyzed separately (Figure 4C, 4D), right eye and mouth fixations elicited larger amplitudes compared to nose fixation ($ps \le .001$), whereas the N170 amplitude for mouth fixation did not differ from amplitudes elicited by either eye ($ps \ge .11$; simple effect of Featural Fixation: $F(1.98, 65.43) = 10.94$, $MSE = 4.92$, $p < .001$, $\eta_p^2 = .25$). Thus, the mouth context effect was due to a greater N170 amplitude reduction for the mouth compared to eye fixations within the face context (Figure 4A, 4D).

3.1.2 N170 peak latency.

A robust main effect of Context was due to significantly longer N170 latencies for isolated features relative to fixation on the same features within a face context ($F(1,33) = 443.25$, $MSE = 118.63$, $p < .001$, $\eta_p^2 = .93$), yielding an average delay of approximately 25 ms. This delay varied between features, as revealed by a Context by Featural Fixation interaction ($F(3,99) = 70.25$, $MSE = 16.31$, $p < .001$, $\eta_p^2 = .68$; Figure 4A and 4F).

N170 context difference scores (isolated feature latency – face latency) were calculated for each featural fixation. Analysis revealed that nose fixation yielded the largest context difference of 33 ms ($ps < .001$), followed by the eyes (which did not differ from each other), followed in turn by the mouth (left eye: $p = .08$; right eye: $p = .02$). When faces were analyzed separately, nose fixation elicited the shortest N170 latency ($ps < .001$), while mouth fixation yielded the longest latency (nose: $p < .001$; right eye: $p = .004$; left eye: $p = .13$; simple effect of Featural Fixation:

---

[7] The same was found when the analysis on face conditions included the nasion fixation. The N170 amplitude for nasion fixation was also significantly smaller than for eye fixations ($p < .001$) but not significantly different than that recorded for nose or mouth fixations ($ps \ge .68$). See Supplementary Figure.

$F(3,99) = 46.38$, $MSE = 13.18$, $p < .001$, $\eta_p^2 = .58$)[8]. Alternatively, when isolated features were

analyzed separately, nose fixation yielded the longest latency relative to all other features ($ps <$

.001), which did not differ from each other ($ps = 1.00$; Figure 4C, 4F). Thus, nose fixation yielded

the largest latency difference score (Figure 4G) due to both shorter latencies when presented within

the context of a face, and longer latencies when presented in isolation.

A main effect of Hemisphere was also found, with faster N170 latencies in the right

hemisphere compared to the left ($F(1,33) = 6.52$, $MSE = 114.95$, $p = .02$, $\eta_p^2 = .17$). Furthermore,

the Hemisphere by Featural Fixation interaction ($F(3,99) = 10.19$, $MSE = 16.34$, $p < .001$, $\eta_p^2 =$

.24) was qualified by the Hemisphere by Context by Featural Fixation interaction ($F(3,99) = 22.93$,

$MSE = 18.09$, $p < .001$, $\eta_p^2 = .41$). Follow-up Bonferroni-corrected paired t-tests (significance level

at $p \leq .013$) performed on context latency difference scores (isolated feature latency – face latency)

revealed larger differences in the hemisphere ipsilateral to eye fixation compared to the

contralateral hemisphere ($ps < .001$; Figure 4G). Alternatively, no significant hemispheric

differences were observed for nose or mouth fixation latency difference scores ($ps = .30$ and .13,

respectively).

---

[8] When nasion fixation was included in this analysis, the fastest latency was found for both nasion and nose fixations, which did not differ significantly from each other ($p = 1.00$), and were both shorter than all other fixation locations ($ps < .001$). See Supplementary Figure.
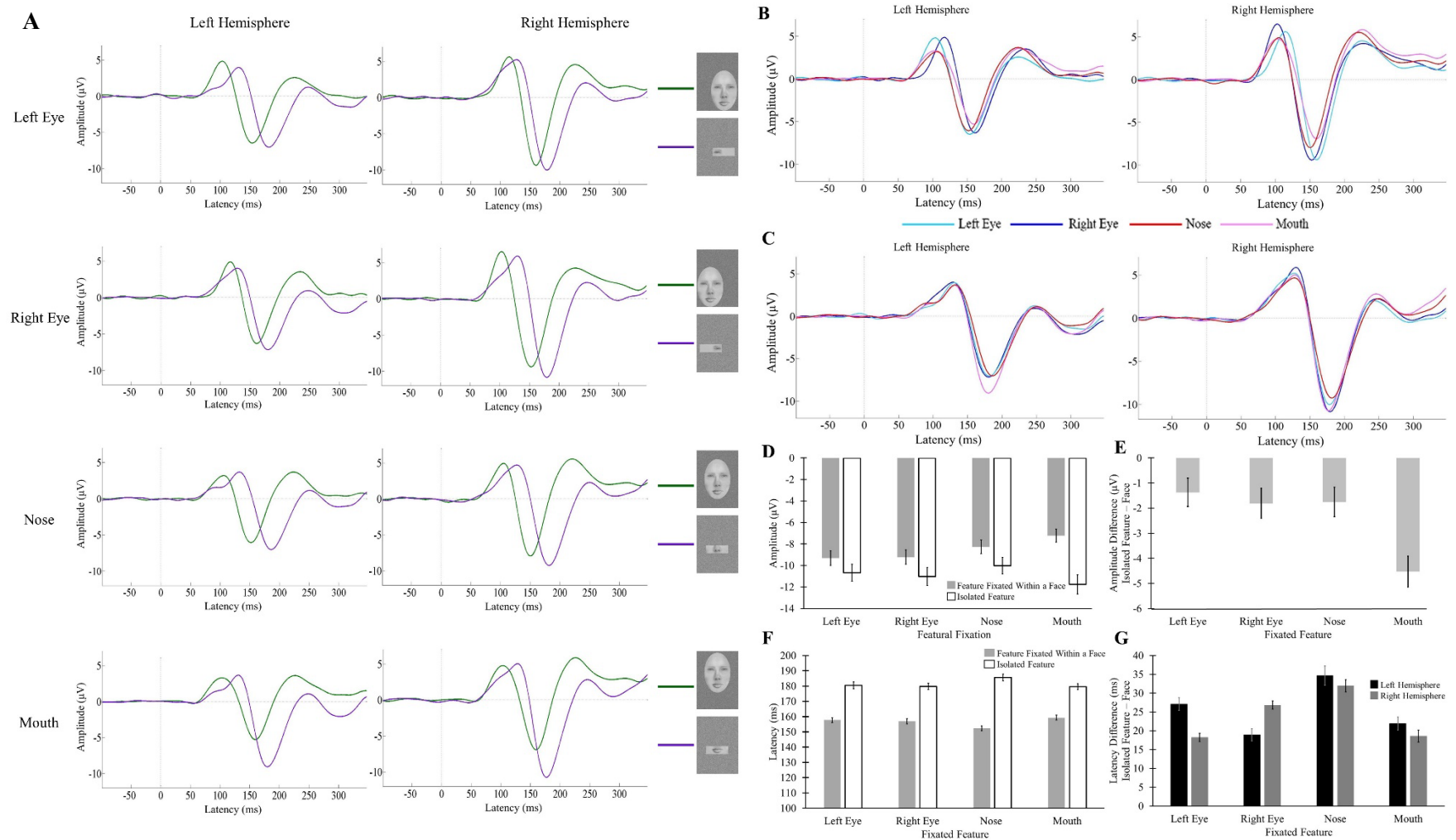
Figure 4: Peak N170 amplitudes and latencies for isolated features and features fixated within a face, averaged across the electrodes at which the N170 was maximal for each participant. All error bars signify ± 1 standard error (SE). *Panel A:* N170 ERP waveforms for isolated features (purple) and features fixated within a face context (green). *Panel B:* N170 ERP waveforms for features fixated within a face. Note the enhanced N170 amplitudes for left and right eye fixations, especially in the right hemisphere. *Panel C:* N170 ERP waveforms for isolated features. Note the absence

of an isolated single eye sensitivity. *Panel D:* Mean N170 peak amplitudes for each fixated feature comparing isolated and face conditions (averaged across hemispheres). *Panel E:* Mean N170 amplitude difference scores (isolated feature – face) for all fixated features (averaged across hemispheres). Note the largest difference for mouth fixation. *Panel F:* Mean N170 peak latencies or all fixated features, comparing isolated and face conditions. *Panel G:* Mean N170 latency difference scores (isolated feature – face) for all fixated features. Note the larger latency difference for nose fixation relative to all other featural fixations, and the hemispheric effect for eyes.

3.1.3 Results summary.

As predicted, the N170 was larger and delayed for isolated features compared to the same features fixated within a face. This context effect was further moderated by the feature fixated, although different patterns emerged for amplitude and latency measures. For N170 peak amplitude, we found that the mouth elicited the largest context effect, which was driven by an amplitude reduction for mouth fixation compared to eye fixations within the face context, while single isolated eyes and mouths elicited similar N170 amplitudes. In terms of peak N170 latency, the largest context effect was found for the nose, driven by both shorter latencies when presented within the context of a face and longer latencies when presented in isolation. We also replicated the strong eye sensitivity within faces, despite the internal facial features being equalized for pixel intensity and RMS contrast. However, an eye sensitivity was not seen when only one isolated eye was presented, compared to an isolated mouth or nose.

3.2. Impact of Eye Context: Single Eye vs. Eye Region

3.2.1 Left and right eye fixations.

Peak N170 amplitude and latency modulations to single isolated eyes, eye regions, and faces, when fixation was enforced on an eye, were analyzed separately using repeated measure ANOVAs with Eye Fixation (2: left eye, right eye), Eye Context (3: single isolated eye, eye region,

face), and Hemisphere (2: left, right) as within-subject factors. The following conditions were included: ILE-LE, IRE-RE, IEyes-LE, IEyes-RE, FF-LE, and FF-RE (Figure 1A). This analysis examined the N170 eye sensitivity, with the prediction that eye regions would elicit larger and faster N170 responses relative to single eyes, due to cumulative coding of foveal and parafoveal eye information.

3.2.1.1 N170 peak amplitude.

A main effect of Hemisphere was driven by larger N170 peak amplitudes in the right hemisphere compared to the left ($F(1,33) = 16.92$, $MSE = 76.39$, $p < .001$, $\eta_p^2 = .34$); this hemispheric difference was more pronounced for right eye than left eye fixation (Hemisphere by Eye Fixation interaction: $F(1,33) = 50.66$, $MSE = 2.81$, $p < .001$, $\eta_p^2 = .35$). Critically, a main effect of Eye Context was due to faces eliciting smaller N170 amplitudes than single eyes, which in turn elicited smaller amplitudes than eye regions ($F(1.45, 47.70) = 23.07$, $MSE = 9.47$, $p < .001$, $\eta_p^2 = .41$; Figure 5A-C). However, this effect was qualified by Hemisphere by Eye Context ($F(1.42, 46.87) = 4.29$, $MSE = 4.80$, $p = .031$, $\eta_p^2 = .12$) and Hemisphere by Eye Fixation by Eye Context ($F(2,66) = 10.85$, $MSE = 1.76$, $p < .001$, $\eta_p^2 = .25$) interactions.

To quantify the three-way interaction, Hemisphere by Eye Context ANOVAs were conducted separately for each eye fixation. This interaction was significant for right eye fixation ($F(1.64, 54.08) = 12.59$, $MSE = 3.21$, $p < .001$, $\eta_p^2 = .28$), with faces eliciting consistently smaller N170 amplitudes in both hemispheres, compared to eye regions and single eyes ($ps \leq .002$). However, eye regions only elicited larger N170 amplitudes compared to single eyes in the right hemisphere ($p = .003$; Figure 5C). In contrast, the interaction was not significant for left eye fixation ($p = .76$) which showed only the main effect of Eye Context. Thus, eye regions with left eye fixation elicited larger amplitudes than single eyes in both hemispheres, whereas enhanced

amplitudes for eye regions compared to single eyes with right eye fixation were only found in the right hemisphere (Figure 5A-C).

3.2.1.2 N170 peak latency.

A robust main effect of Eye Context was due to significantly faster N170 latencies for faces, followed by eye regions, followed in turn by single eyes ($F(1.58, 52.27) = 377.17$, $MSE = 59.17$, $p < .001$, $\eta_p^2 = .92$; paired comparisons $ps < .001$; Figure 5A, 5B, 5D). Importantly, this effect was largest in the hemisphere ipsilateral to eye fixation, as revealed by significant Hemisphere by Eye Fixation ($F(1,33) = 73.15$, $MSE = 12.93$, $p < .001$, $\eta_p^2 = .69$) and Hemisphere by Eye Fixation by Eye Context ($F(2,66) = 49.77$, $MSE = 13.30$, $p < .001$, $\eta_p^2 = .60$) interactions. As seen in Figure 5D, the N170 latency difference between the full face, eye region, and single eyes was observed for both left and right eye fixations, but was most pronounced in the hemisphere ipsilateral to eye fixation, due to the presence of the other eye in the opposing visual field.
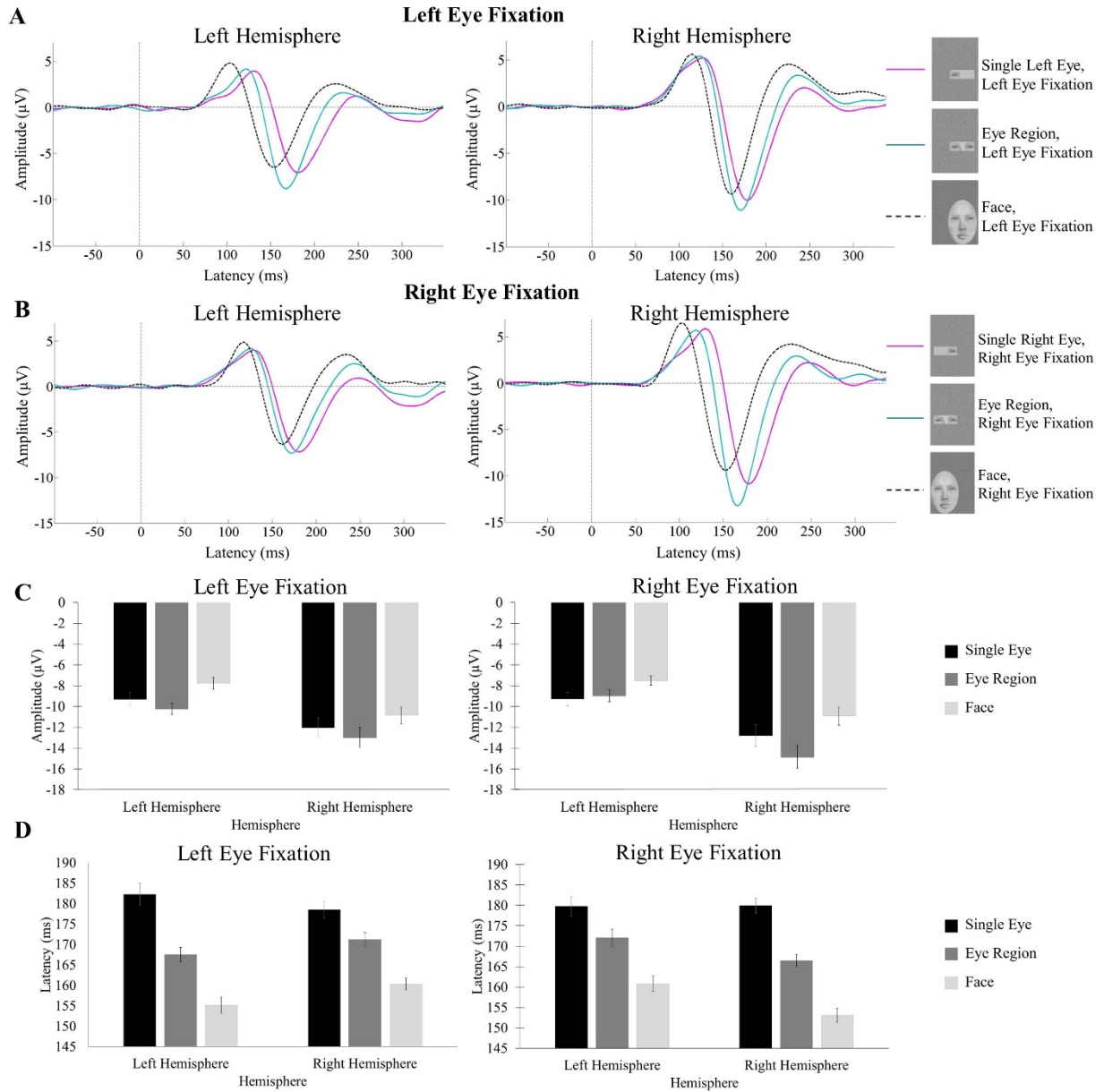
Figure 5. Peak N170 amplitudes and latencies for face and isolated eye conditions with left and right eye fixations, averaged across the electrodes at which the N170 was maximal for each participant. All error bars signify ± 1 SE. *Panel A:* Group ERP waveforms for face and isolated eye conditions with left eye fixation. *Panel B:* Group ERP waveforms for face and isolated eye conditions with right eye fixation. *Panel C*: Mean N170 amplitudes for face and isolated eye conditions with left and right eye fixations in each hemisphere. Note the largest N170 amplitudes for eye regions, followed by single eyes and then faces. *Panel D:* Mean N170 latencies for face and isolated eye conditions with left and right eye fixations in each hemisphere. Note the shorter N170 latencies to faces, followed by eye regions, then single eyes; an effect enhanced in the hemisphere ipsilateral to eye fixation (i.e., contralateral to parafoveal eye/face information).

3.2.2. Nasion fixation.

Peak N170 amplitude and latency modulations to single isolated eyes, eye regions, and faces, when fixation was enforced on the nasion (and thus eyes were in parafovea), were analyzed separately using repeated measure ANOVAs with Eye Context (4: single left eye, single right eye, eye region, face), and Hemisphere (2: left, right) as within-subject factors. The following conditions were included: ILE-Na. IRE-Na, IEyes-Na, and FF-Na (Figure 1A). Similar to that predicted for eye fixations, it was expected that eye regions would elicit larger and faster N170 responses compared to single eyes.

3.2.2.1 N170 peak amplitude.

N170 amplitudes were larger in the right hemisphere compared to the left hemisphere ($F(1,33) = 13.07$, $MSE = 33.72$, $p = .001$, $\eta_p^2 = .28$). A main effect of Eye Context was also significant ($F(2.14, 70.62) = 22.87$, $MSE = 7.32$, $p < .001$, $\eta_p^2 = .41$), but was further qualified by a Hemisphere by Eye Context interaction ($F(2.34, 77.07) = 8.91$, $MSE = 5.83$, $p < .001$, $\eta_p^2 = .21$). As shown in Figure 6A-B, eye regions yielded N170 amplitudes that were consistently larger than faces and single eyes in both hemispheres ($ps \leq .001$ for all comparisons except right eye fixation in the right hemisphere where $p = .02$). Within the left hemisphere, single left and right eyes did not differ from each other ($p = .19$) nor from faces ($p = 1.00$). However, within the right hemisphere, although single left and right eyes did not differ from faces ($ps \geq .06$), single left eyes elicited larger N170 amplitudes compared to single right eyes ($p = .001$). Bonferroni-corrected paired t-tests (significance level at $p \leq .013$) further revealed consistently larger N170 amplitudes in the right than in the left hemisphere for single left eye ($p < .001$), eye region ($p = .001$), and face ($p = .009$) conditions, but no hemispheric differences for single right eyes ($p = .43$).

3.2.2.2 N170 peak latency.

A main effect of Eye Context was significant, such that faces yielded the fastest latencies ($ps < .001$), followed by eye regions, which were in turn faster than single eyes ($ps < .001$; $F(3,99) = 115.74$, $MSE = 88.93$, $p < .001$, $\eta_p^2 = .78$). Single right eyes also yielded slower N170 peak latencies relative to single left eyes ($p = .02$). This effect was qualified further by a significant Hemisphere by Eye Context interaction ($F(2.08, 68.48) = 13.40$, $MSE = 135.40$, $p < .001$, $\eta_p^2 = .29$), with faster N170 latencies in the left compared to the right hemisphere for single right eyes ($p = .005$), and faster N170 latencies in the right compared to the left hemisphere for single left eyes ($p = .001$; Figure 6A and 6C). Alternatively, face and eye region conditions elicited similar latencies in both hemispheres ($ps \geq .29$).
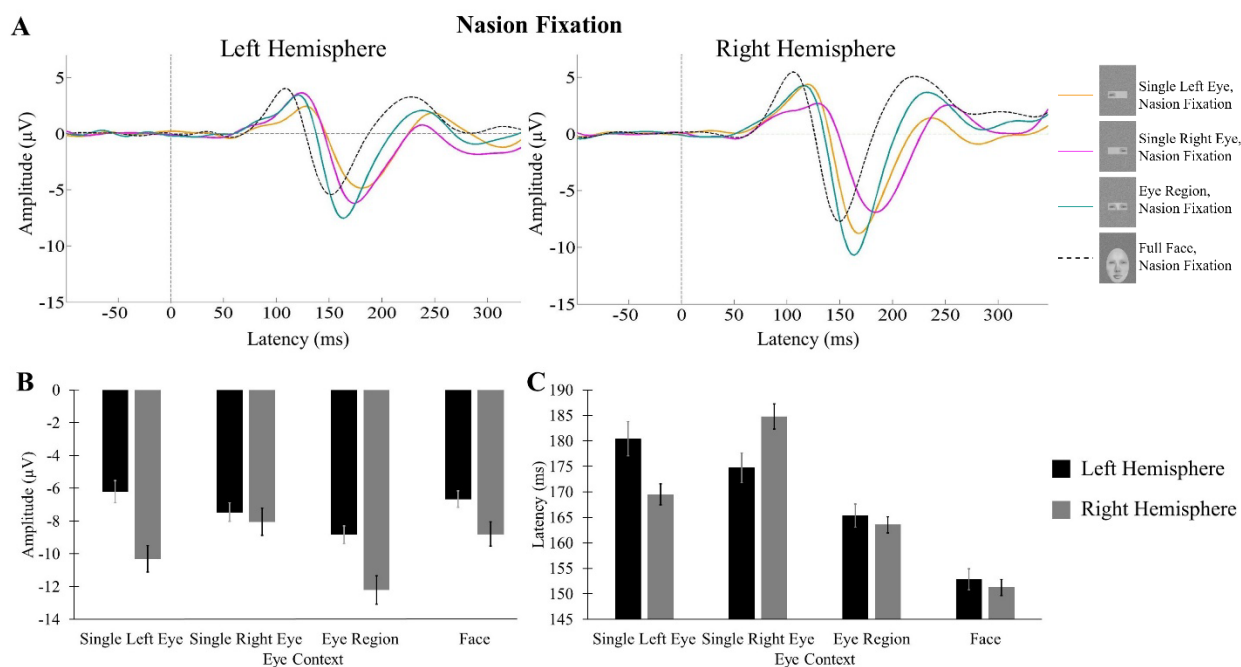

Figure 6: Peak N170 amplitudes and latencies for face and isolated eye conditions with nasion fixation, averaged across the electrodes at which the N170 was maximal for each participant. All error bars signify ± 1 SE. *Panel A:* Group ERP waveforms for face and isolated eye conditions with nasion fixation. *Panel B:* Mean N170 amplitudes for face and isolated eye conditions with nasion fixation in both hemispheres. Note the larger N170 amplitudes for eye regions relative to single eyes, and the lack of right hemispheric lateralization for the single right eye. *Panel C:* Mean N170 latencies for isolated eye conditions with nasion fixation in both hemispheres. Note the

slower N170 latencies for single eyes compared to eye regions, particularly in the hemisphere ipsilateral to parafoveal eye information.

3.2.3. Results summary.

Isolated eye regions elicited enhanced N170 peak amplitudes relative to faces and single isolated eyes, irrespective of left eye, right eye, or nasion fixation. This enhanced neural coding of eye regions compared to single eyes was seen in both hemispheres when fixation was enforced on the left eye (with the right eye in the right visual field) or nasion, but was only seen in the right hemisphere when fixation was enforced on the right eye (with the left eye in the left visual field). Eye regions also yielded intermediate N170 peak latencies relative to those observed for faces and single isolated eyes, an effect that was enhanced in the hemisphere ipsilateral to eye fixation (i.e., contralateral to the other, non-fixated eye).

4. DISCUSSION

The current study sought to further illuminate the neural response to the eyes in early face perception. In a gaze-contingent ERP paradigm, N170 responses were recorded to isolated facial features (left eye, right eye, nose, and mouth) and the same features fixated within a face context. Direct comparison of featural responses across contexts permitted further examination of the N170 eye sensitivity, and of the inhibition mechanism proposed by the LIFTED model of early face perception (Nemrodov et al., 2014). N170 responses to single isolated eyes, a condition not tested before, and to the classically-used eye region, were also compared to refine our understanding of what makes the eyes so special.

Consistent with previous reports, we replicated the N170 sensitivity to eye fixations within a face, with maximal N170 peak amplitude, when fixation was enforced on an eye, compared to nose and mouth fixations (de Lissa et al., 2014; Itier & Preston, 2018; Neath & Itier, 2015; Neath-

Tavares & Itier, 2016; Nemrodov et al., 2014). It is also apparent that this eye sensitivity is specific to direct fixation on an eye, rather than simply looking within the general eye region, as nasion fixation also elicited attenuated responses compared to left and right eye fixations (Supplementary Figure). Because mean pixel intensity and RMS contrast were equalized at the local (featural) level in the present study, we conclude that this eye sensitivity is the result of higher-order perceptual mechanisms rather than the mere effect of these local low-level visual differences.

The main goal of the present study was to re-evaluate the early visual responses to facial features by comparing gaze-contingent N170 responses to isolated features and fixation to the same features within the context of a face. A robust context effect was found, whereby N170 responses were larger and delayed for all isolated features, compared to enforced fixation on the same features within a face. Previous studies have reported larger and delayed responses for isolated eye regions compared to faces (Bentin et al., 1996; Taylor, Edmonds, et al., 2001a; Taylor, Itier, et al., 2001b; Cauquil et al., 2000; Itier et al., 2006, 2007, 2011; Nemrodov et al., 2011; Shibata et al., 2002). However, the few studies that included isolated mouth (lips) and/or nose features reported delayed but *smaller* N170 amplitudes (Bentin et al., 1996; Taylor, Itier, et al., 2001b) or no significant differences (Nemrodov & Itier, 2011) for those features compared to eye regions or full faces. This discrepancy with the current findings is likely due to the lack of stimulus and gaze position control in these earlier reports. Indeed, previous studies did not always equalize pixel intensity and contrast at the global (image) level between faces and isolated features, nor were eye movements monitored, making it unclear what information participants were focusing on during the ERP recordings. In fact, previous studies have primarily used a single, centrally-presented face category as the default comparison to isolated features, with fixation presumably in the centre, around the nose or nasion. Here, we show that when each feature is compared to its

direct equivalent within a face context, amplitudes are larger for *all* isolated features, including single eyes, a condition not tested before. Because the same feature was situated in fovea in both context conditions, any N170 modulations are attributable to the presence of parafoveal information in the face context. Thus, the presence of parafoveal features decreased the overall N170 amplitude, a pattern in line with the neural inhibition mechanism proposed by the LIFTED model (Nemrodov et al., 2014), according to which the neurons coding for parafoveal information inhibit the neurons coding for foveal information when the normal upright face configuration is detected.

The present results also show that when fixation is controlled within the face, the N170 amplitude difference between an isolated feature and its equivalent fixation within a whole face (i.e. the context effect), is not uniform, but instead varies depending on which feature is fixated. Specifically, mouth fixation generated the largest N170 amplitude difference (an astonishing average of 4.53 µV ; Figure 4E), which was entirely due to a sharp decrease in amplitude when the mouth was fixated within the face compared to when presented in isolation (isolated single eyes and isolated mouths elicited similar amplitudes). Within the LIFTED model framework, this amplitude decrease in the face context can be interpreted as a large inhibition of the neurons coding for the mouth in fovea by the neurons coding for the other features situated in parafovea. In contrast, nose, left eye, and right eye fixations elicited smaller amplitude context effects (<2µV) that did not differ significantly from one another.

The smaller amplitude context effect for the eyes compared to the mouth was expected, as the eyes should only be partially inhibited compared to other features. Specifically, the LIFTED model proposes that the inhibition mechanism likely depends on the retinotopic distances between features, and thus on face size, with less inhibition from features situated farther away from fovea

than from features situated closer to fovea. The eye sensitivity is expected to persist (and thus inhibition mechanisms should be weaker for the eyes) because, when the eyes are fixated, the inhibition signals from the neurons coding for parafoveal features are, arguably, not strong enough to completely inhibit the activity of the eye detector. As the faces used in the present study were slightly larger than the ones used by Nemrodov et al. (2014), we expected less inhibition of the eyes here than reported previously, and thus an even clearer eye sensitivity. Thus, the eyes were expected to be less inhibited than the mouth or nose. However, although we found less inhibition for the eyes compared to the mouth, this eye inhibition was similar to the inhibition found for the nose, a result not expected.

The inhibition (context effect) for the nose was the result of amplitude variations in both context conditions. Isolated noses elicited slightly smaller amplitudes than isolated mouths and eyes. However, within the context of a face, nose fixation elicited an N170 amplitude intermediate between that elicited by mouth and eye fixations. Overall, within the face context, N170 amplitudes were largest for eye fixations, followed by nose fixation, and smallest for mouth fixation. This pattern possibly reflects a particular sensitivity of extrastriate visual areas to the location of eyes within the visual field, with largest N170 responses when an eye is in fovea, and decreasing amplitude with increasing eccentricity of the eyes from fovea (nose and mouth fixations were situated 3.52° and 5.31°, respectively, from the centre of each eye, see Figure 1). This coding of distances to the eyes is in line with the LIFTED model's proposal that the eyes serve as anchoring points from which the position of the other facial features is coded based on a face template (Nemrodov & Itier, 2011; Nemrodov et al., 2014). However, this particular gradient of amplitude responses, which was not found in previous studies (Itier & Preston, 2018; Neath &

Itier, 2015; Neath-Tavares & Itier, 2016; Nemrodov et al., 2014), does not completely fit with the current LIFTED model.

According to the LIFTED model, there are two mechanisms at play for upright faces during the N170 timing: the activity of an eye detector (only partially inhibited), and holistic processing of the rest of the face. Strict holistic processing, by definition, entails that features are perceptually equal and integrated into an *indecomposable* perceptual whole. Fixating on the nose or mouth (or anywhere else on the face), in that view, should elicit similar N170 amplitudes. In the inhibition logic of the LIFTED model, the nose should be inhibited more so than the eyes, but to the same degree as the mouth. The larger N170 amplitude observed here, for nose versus mouth fixations within the face, points at possible differential inhibition mechanisms related to the nature of the fixated feature (which is inherently linked to its position within the face), rather than its distance to other parafoveal features (see also Itier & Preston, 2018). The present results thus suggest that the process might not be strictly holistic but rather involves a complex integration of features, which varies with fixation location. This integration process, more malleable than the rigid "perceptual snapshot" that a strict holistic processing assumes, might be influenced by factors such as individual preferences in fixation positions. Indeed, individual differences in attention to internal facial features have been reported at the behavioural level (Haig, 1985; Kanan et al., 2015; Mehoudar et al., 2014; Miellet et al., 2011), and may contribute to the present amplitude gradient, an idea that should be investigated in future reports.

Alternatively, N170 latency context effects were different than those seen for amplitude. Although latencies were shorter for all features fixated within faces compared to isolated features (Figure 4A and 4F), nose fixation elicited the largest latency decrease (33 ms), whereas fixation on the eyes and mouth yielded comparatively smaller latency reductions (20-23 ms). This latency

context effect was due to both longer latencies for isolated noses and shorter latencies for nose fixation within the face. The latter finding within a face context was also reported by Nemrodov et al. (2014). The authors interpreted this as being due to maximal holistic face processing, such that fixation on the nose (which is situated in a face's relative centre of mass) permits easier viewing of other features within parafovea, and thus initiates fastest integration of the facial features. Taken together, the present findings suggest that N170 amplitude and latency measures may be differentially sensitive to feature integration into the face percept, an idea that will have to be investigated further by future studies.

Bentin et al. (1996) originally interpreted the larger N170 response to the eye region compared to isolated noses and mouths as being in line with the feature saliency theory proposed by Shepherd et al. (1981), whereby the eyes would be more salient than the mouth, itself more salient than the nose. Although the present results showed larger N170 amplitudes to isolated single eyes and mouths compared to isolated noses, they also showed no significant difference between isolated mouths and single eyes, going against a true eye sensitivity or saliency at the individual feature level. Within a feature saliency framework, the eyes are considered the most diagnostic features for various face processes (e.g. gender discrimination, identity and emotion recognition), followed by the mouth, followed in turn by the nose (Hills, Ross, & Lewis, 2011; Schyns, Petro, & Smith, 2007; Shepherd et al., 1981). However, it should be noted that these studies most commonly assessed the eye region, and thus, the specific contribution of each individual eye has always been unclear. Here, we demonstrate that there is nothing truly special about a single eye feature in terms of the N170 amplitude response. Rather, the saliency of the eyes seems to be driven by the presence of two symmetrical eyes (eye region).

Indeed, a specific eye sensitivity emerged when the eye region was presented. Consistent with past research (Bentin et al., 1996; Itier et al., 2007, 2006, 2011; Kloth et al., 2013; Taylor, Edmonds, et al., 2001a; Taylor, Itier, et al., 2001b), isolated eye regions elicited larger N170 amplitudes compared to faces, even with fixation enforced to the eyes of the face. In addition, and consistent with the LIFTED model, eye regions also elicited larger N170 amplitudes than isolated single eyes, an effect seen irrespective of nasion or eye fixation. The LIFTED model predicts that eye regions do not activate the neural inhibition mechanism due to the disruption of the face configuration, resulting in each eye contributing, in a combined fashion, to the overall neural response. Moreover, the present findings suggest that this eye region sensitivity is not due to sensory competition between two simultaneously presented eyes, but rather to the eye region being processed as a single percept. According to the sensory competition theory, the neural representation of a visual image is down-regulated when two images of the same category are presented simultaneously, compared to the presentation of one image on its own (for review see Beck & Kastner, 2009). Faces seem to follow this rule, presumably because visual information provided by each face is unique (i.e., different person identities, possibly with different gaze directions or emotional expressions). Indeed, it has been shown that when two or more faces are presented simultaneously, the N170 is attenuated (Jacques & Rossion, 2004, 2007) and a reduced BOLD response is observed bilaterally within the fusiform face area (FFA) – one of the main brain nodes of the face perception network (Haxby, Hoffman, & Gobbini, 2000) – as well as within right lateral occipital cortex (rLOC) regions (Nagy, Greenlee, & Kovács, 2011). Importantly, however, the occipital face area (OFA) – another main brain node of the face network thought to be involved in feature perception – does not seem to be affected by sensory competition (Nagy et al., 2011), suggesting that features, unlike whole faces, may not be subject to sensory competition. Our results

are in line with this view, as we found the opposite of what sensory competition would predict for the eyes (i.e., larger N170 responses for the eye region compared to one single eye). Because two eyes provide complimentary and redundant social information (e.g., same gaze direction or emotional expression), they might be perceptually glued together into a single percept, rather than compete for representation, in order to reduce perceptual redundancy and optimize social relevance.

Eye regions also elicited N170 peak latencies that fell intermediate to N170 latencies recorded to single eyes and to fixation on eyes within a face. Specifically, relative to fixation within a face, single isolated eyes were delayed by as much as 27 ms in the hemisphere ipsilateral to eye fixation, (Figures 4G and 5D), and by 18 ms in the hemisphere contralateral to eye fixation, while the delay for eye regions compared to faces was only about 10-13ms (depending on the eye fixated and hemisphere), a range in line with past reports (Bentin et al., 1996; Itier et al., 2007, 2006, 2011; Kloth et al., 2013; Taylor, Edmonds, et al., 2001a). It should be noted that the LIFTED model made clear predictions for amplitude but was less clear for latencies. This intermediate N170 latency for eye regions compared to full faces and single isolated eyes may reflect an intermediate level of processing where the presence of the second eye in parafovea boosts the processing speed, in a manner similar to how the presence of (both) eyes in a face boosts the latency response compared to eyeless faces (Eimer, 1998; Itier et al., 2007, 2011; Kloth et al., 2013; Nemrodov et al., 2014).

Notably, the present data are also in line with intracranial N200 ERP findings recorded directly from the surface of the ventral (vOTC) and lateral (lOTC) occipitotemporal cortex in human patients (e.g., Engell & McCarthy, 2014; McCarthy, Puce, Belger, & Allison, 1999). In particular, N200s elicited by faces occur earlier and are attenuated relative to N200s elicited by

eye regions, similar to the N170 responses reported here. Furthermore, McCarthy and colleagues (1999) identified vOTC and lOTC sites where face parts (eye regions, lips, and noses) elicited larger and later N200 responses compared to whole faces, consistent with the present scalp-recorded N170 data. More recently, Engell and McCarthy (2014) reported that the majority of the intracranial sites were eye region-selective[9], rather than face-selective, with many truly eye-specific[10] sites. Thus, the OTC seems to be more sensitive to eye information than to the whole face, in agreement with recent cell recordings in non-human primates (Freiwald, Tsao, & Livingstone, 2009; Meyers, Borzello, Freiwald, & Tsao, 2015). Interestingly, the majority of eye region-specific intracranial sites were localized to the posterior lOTC, which corresponds well with the occipital face area. Converging fMRI evidence further suggests that the OFA is most strongly activated during the presentation of single facial features (left eye, nose, and mouth) and eye regions, with monotonic response decreases for three-feature or full face stimuli (Arcurio, Gold, & James, 2012), in line with the LIFTED model's proposed inhibition mechanism. It is noteworthy, however, that Arcurio et al. (2012) reported similar degrees of OFA activation for eye regions and single left eyes[11], which is in contrast to the larger N170 response for eye regions compared to single eyes or mouths observed in the present study. Taken together, these intracranial and fMRI findings suggest that brain generators situated on the posterior lOTC (possibly corresponding to the OFA) could significantly contribute to the scalp-recorded N170, and explain the present findings.

---

[9] Selectivity was defined as a response twice as large for eye regions compared to a control flower category (Engell & McCarthy, 2014).

[10] Specificity was defined as a response twice as large for eye regions compared to faces, bodies, and a control flower category (Engell & McCarthy, 2014).

[11] It should be noted that Arcurio et al. (2012) refer to this condition as "right eye" due to different points of reference, but this corresponds to our observer's left eye condition, so we refer to it as left eye here for consistency.

Within an evolutionary context, the refined ability to perceive and detect two symmetrical eyes is arguably most adaptive (Emery, 2000; Kobayashi & Kohshima, 1997). The perception of a single eye-like stimulus is not particularly diagnostic of the presence of an animal or person, as many objects contain similar characteristics (i.e., small, round shape; high local contrast). Alternatively, the presence of two horizontally symmetrical eye-like stimuli are more likely to be indicative of a pair of eyes, thereby indicating the presence of another person or predator. Therefore, the easy and rapid detection of an eye region (rather than a single eye) can lead to more effective identification of potential threat or social interaction, and it thus makes sense that the neural response reflects the activity of an eye *region* detector rather than a single eye detector. Overall, the present scalp-ERP findings, together with recent intracranial ERPs and fMRI results, provide compelling evidence supporting the existence of an eye region detector at the neural level, with potentially different underlying neural generators for face and feature information.

5. CONCLUSION

The present study confirms the sensitivity of the N170 ERP component to the eyes within a face, and further suggests that this sensitivity is not merely driven by low-level factors (pixel intensity and contrast), but rather reflects higher level processes. Maximal sensitivity was seen to the presence of two eyes (eye region) presented outside a face context, while response to a single eye did not differ strongly from that to other isolated features. These findings strongly suggest that early face perception mechanisms are not purely holistic, contrary to previous assumption. In particular, the current results suggest that early face perception may be better characterized by a complex integration of featural information that vary with fixation location, rather than by a rigid holistic process. While both amplitude and latency capture these neural mechanisms, they seem to be maximally sensitive to different fixation locations. Overall, the present findings are largely in

line with the LIFTED model's view according to which the N170 ERP component reflects two combined processes. On one hand, the N170 seems to reflect the activity of an eye region detector maximally sensitive to the detection of two symmetrical eyes in the environment, a step necessary to anchor the face percept according to a pre-existing face template representation. On the other hand, the N170 also seems to reflect the output of the neural inhibition mechanism proposed to facilitate feature integration into a perceptual whole face (i.e., the inhibition of foveal content by parafoveal face information), a process that seems non-uniform but rather varies with fixation location.

Acknowledgements

Declarations of Interest: none

References

Arcurio, L. R., Gold, J. M., & James, T. W. (2012). The response of face-selective cortex with single face parts and part combinations. *Neuropsychologia*, *50*(10), 2454–2459. doi: 10.1016/j.neuropsychologia.2012.06.016

Beck, D.M. & Kastner, S. (2009). Top-down and bottom-up mechanisms in biasing competition in the human brain. *Vision Research, 49*(10), 1154-1165. doi: 10.1016/j.visres.2008.07.012

Bentin, S., Allison, T., Puce, A., Perez, E., & McCarthy, G. (1996). Electrophysiological studies of face perception in humans. *Journal of Cognitive Neuroscience*, *8*(6), 551–565. doi: 10.1162/jocn.1996.8.6.551

Birmingham, E., B ischof, W.F., & Kingstone, A. (2008a). Gaze selection in complex social scenes. *Visual Cognition, 16*(2/3), 341-355. doi: 10.1080/13506280701434532

Birmingham, E., Bischof, W.F., & Kingstone, A. (2008b). Social attention and real-world scenes: The roles of action, competition and social content. *Quarterly Journal of Experimental Psychology, 61*(7), 986-998. doi: 10.1080/17470210701410375

Cauquil, A.S., Edmonds, G.E., & Taylor, M.J. (2000). Is the face-sensitive N170 the only ERP not affected by selective attention? *NeuroReport, 11*(10), 2167-2171. doi: 10.1097/00001756-200007140-00021

de Lissa, P., McArthur, G., Hawelka, S., Palermo, R., Mahajan, Y., & Hutzler, F. (2014). Fixation location on upright and inverted faces modulates the N170. *Neuropsychologia*, *57*, 1–11. doi: 10.1016/j.neuropsychologia.2014.02.006

Delorme, A. & Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, *134*(1), 9–21. doi: 10.1016/j.jneumeth.2003.10.009

Eimer, M. (1998). Does the face-specific N170 component reflect the activity of a specialized eye processor? *NeuroReport, 9*(13), 2945-2948. doi: 10.1097/00001756-199809140-00005

Eimer, M. (2000a). Event-related brain potentials distinguish processing stages involved in face perception and recognition. *Clinical Neurophysiology*, *111*(4), 694–705. doi: 10.1016/S1388-2457(99)00285-0

Eimer, M. (2000b). The face-specific N170 component reflects late stages in the structural encoding of faces. *Neuroreport*, *11*(10), 2319–2324. doi: 10.1097/00001756-200007140-00050

Emery, N. J. (2000). The eyes have it: the neuroethology, function and evolution of social gaze. *Neuroscience and Biobehavioral Reviews*, *24*(6), 581–604. doi: 10.1016/S0149-7634(00)00025-7

Engell, A.D. & McCarthy, G. (2014). Face, eye, and body selective responses in fusiform gyrus and adjacent cortex: An intracranial EEG study. *Frontiers in Human Neuroscience, 8*(642), 1-12. doi: 10.3389/fnhum.2014.00642

Freiwald, W. A., Tsao, D. Y., & Livingstone, M. S. (2009). A face feature space in the macaque temporal lobe. *Nature Neuroscience*, *12*(9), 1187–1196. doi: 10.1038/nn.2363

George, N., Evans, J., Fiori, N., Davidoff, J., & Renault, B. (1996). Brain events related to normal and moderately scrambled faces. *Cognitive Brain Research*, *4*(2), 65–76. doi: 10.1016/0926-6410(95)00045-3

Haig, N.D. (1985). How faces differ - A new comparative technique. *Perception, 14*(5), 601-615. doi: 10.1068/p140601

Haxby, J. V., Hoffman, E. A., & Gobbini, M. I. (2000). The distributed human neural system for face perception. *Trends in Cognitive Sciences*, *4*(6), 223–233. doi: 10.1016/S1364-6613(00)01482-0

Henderson, J.M., Williams, Carrick, C., & Falk, R.J. (2005). Eye movements are functional during face learning. *Memory & Cognition, 33*(1), 98-106. doi: 10.3758/BF03195300

Hills, P.J., Cooper, R.E., & Pake, J.M. (2013). First fixations in face processing: The more diagnostic they are the smaller the face-inversion effect. *Acta Psychologia, 142*(2), 211-219. doi: 10.1016/j.actpsy.2012.11.013

Hills, P. J., Ross, D. A., & Lewis, M. B. (2011). Attention misplaced: the role of diagnostic features in the face-inversion effect. *Journal of Experimental Psychology: Human Perception and Performance*, *37*(5), 1396–1406. doi: 10.1037/a0024247

Itier, R.J. (2015). Attention to eyes in face perception. In Fawcett, J.M., Risko, E.F., & Kingstone, A. (Eds.), *The Handbook of Attention* (pp. 369-387). Cambridge, MA: MIT Press.

Itier, R. J., Alain, C., Sedore, K., & McIntosh, A. R. (2007). Early face processing specificity: It's in the eyes! *Journal of Cognitive Neuroscience*, *19*(11), 1815–1826. doi: 10.1162/jocn.2007.19.11.1815

Itier, R.J. & Batty, M. (2009). Neural bases of eye and gaze processing: The core of social cognition. *Neuroscience and Biobehavioral Reviews. 33*(6), 843-863. doi: 10.1016/j.neubiorev.2009.02.004

Itier, R. J., Latinus, M., & Taylor, M. J. (2006). Face, eye and object early processing: what is the face specificity? *NeuroImage*, *29*(2), 667–676. doi: 10.1016/j.neuroimage.2005.07.041

Itier, R.J., & Preston, F.F. (2018). Increased eye sensitivity in mouthless faces: In support of the LIFTED model of early face perception. *Brain Topography*. doi: 10.1007/s10548-018-0663-6

Itier, R. J., Van Roon, P., & Alain, C. (2011). Species sensitivity of early face and eye processing. *NeuroImage, 54*(1), 705–713. doi: 10.1016/j.neuroimage.2010.07.031

Jacques, C., & Rossion, B. (2004). Concurrent processing reveals competition between visual representations of faces. *NeuroReport, 15*(15), 2417-2421. doi: 10.1097/00001756-200410250-00023

Jacques, C., & Rossion, B. (2007). Electrophysiological evidence for temporal dissociation between spatial attention and sensory competition during human face processing. *Cerebral Cortex, 17*(5), 1055-1065. doi: 10.1093/cercor/bhl015

Janik, S.W., Wellens, A.R., Goldberg, M.L., & Dell'osso, L.F. (1978). Eyes as the center of focus in the visual examination of human faces. *Perceptual and Motor Skills, 47*(3 Pt 1), 857-858. doi: 10.2466/pms.1978.47.3.857

Kanan, C., Bseiso, D. N., Ray, N. A., Hsiao, J. H., & Cottrell, G. W. (2015). Humans have idiosyncratic and task-specific scanpaths for judging faces. *Vision Research, 108*, 67–76. doi: 10.1016/j.visres.2015.01.013

Kleinke, C.L. (1986). Gaze and eye contact: A research review. *Psychological Bulletin, 11*(1), 78-100. doi: 10.1016/j.visres.2015.01.013

Kloth, N., Itier, R. J., & Schweinberger, S. R. (2013). Combined effects of inversion and feature removal on N170 responses elicited by faces and car fronts. *Brain and Cognition, 81*(3), 321–328. doi: 10.1016/j.bandc.2013.01.002

Kobayashi, H. & Kohshima, S. (1997). Unique morphology of the human eye. *Nature, 387*, 767–768. doi: 10.1038/42842

McCarthy, G., Puce, A., Belger, A., & Allison, T. (1999). Electrophysiological studies of human face perception II: Response properties of face-specific potentials generated in occipitotemporal cortex. *Cerebral Cortex, 9*(5), 431-444. doi: 10.1093/cercor/9.5.431

McPartland, J., Cheung, C. H., Perszyk, D., & Mayes, L. C. (2010). Face-related ERPs are modulated by point of gaze. *Neuropsychologia, 48*(12), 3657–3660. doi: 10.1016/j.neuropsychologia.2010.07.020

Mehoudar, E., Arizpe, J., Baker, C. I., & Yovel, G. (2014). Faces in the eye of the beholder: Unique and stable eye scanning patterns of individual observers. *Journal of Vision, 14*(7), 1–11. doi: 10.1167/14.7.6

Meyers, E. M., Borzello, M., Freiwald, W. A., & Tsao, D. (2015). Intelligent information loss: The coding of facial identity, head pose, and non-face information in the macaque face patch

system. *The Journal of Neuroscience*, *35*(18), 7069–7081. doi: 10.523/JNEUROSCI.3086-14.2015

Miellet, S., Cladara, R., & Schyns, P. (2011). Local Jekyll and global Hyde: The dual identity of face identification. *Psychological Science, 22*(12), 1518-1526. doi: 10.1177/0956797611424290

Miles, W. R. (1930). Ocular dominance in human adults. *The Journal of General Psychology*, *3*(3), 412–430. doi: 10.1080/00221309.1930.9918218

Nagy, K., Greenlee, M.W., & Kovács, G. (2011). Sensory competition in the face processing areas of the human brain. *PLoS ONE, 6*(9), e24450. doi: 10.1371/journal.pone.0024450

Neath, K. N. & Itier, R. J. (2015). Fixation to features and neural processing of facial expressions in a gender discrimination task. *Brain and Cognition*, *99*, 97–111. doi: 10.1016/j.bandc.2015.05.007

Neath-Tavares, K. N., & Itier, R. J. (2016). Neural processing of fearful and happy facial expressions during emotion-relevant and emotion-irrelevant tasks: A fixation-to-feature approach. *Biological Psychology*, *119*, 122–140. doi: 10.1016/j.biopsycho.2016.07.013

Nemrodov, D., Anderson, T., Preston, F. F., & Itier, R. J. (2014). Early sensitivity for eyes within faces: A new neuronal account of holistic and featural processing. *NeuroImage*, *97*, 81–94. doi: 10.1016/j.neuroimage.2014.04.042

Nemrodov, D. & Itier, R.J. (2011). The role of the eyes in early face processing: A rapid adaptation study of the inversion effect. *British Journal of Psychology, 102*(4), 783-798. doi: 10.1111/j.2044-8295.2011.02033.x

Polyak, S.L. (1941). *The retina.* Chicago: University of Chicago Press.

Rodieck, R.W. (1998). *First steps in seeing.* Sunderland, MA: Sinauer Associates, Inc.

Rossion, B. & Gauthier, I. (2002). How does the brain process upright and inverted faces? *Behavioral and Cognitive Neuroscience Reviews*, *1*(1), 63–75. doi: 10.117/1534582302001001004

Rossion, B. & Jacques, C. (2011). The N170: understanding the time-course of face perception in the human brain. In Kappenman, E.S., & Luck, S.J. (Eds.) *The Oxford Handbook of ERP Components* (pp. 115–142). New York: Oxford University Press.

Rousselet, G. A., Ince, R. A., van Rijsbergen, N. J., & Schyns, P. G. (2014). Eye coding mechanisms in early human face event-related potentials. *Journal of Vision*, *14*(13), 1–24. doi: 10.1167/14.13.7

Schyns, P.G., Bonnar, L., & Gosselin, F. (2002). Show me the features! Understanding recognition from the use of visual information. *Psychological Science, 13*(5), 402-409. doi: 10.1111/1467-9280.00472

Schyns, P. G., Petro, L. S., & Smith, M. L. (2007). Dynamics of visual information integration in the brain for categorizing facial expressions. *Current Biology, 17*(18), 1580–1585. doi: 10.1016/j.cub.2007.08.048

Shepherd, J., Davies, G., & Ellis, H. (1981). Studies of cue saliency. In Davies, G., Ellis, H., & Shepherd, J. (Eds.) *Perceiving and Remembering Faces.* (pp. 105-131).

Shibata, T., Nishijo, H., Tamura, R., Miyamoto, K., Eifuku, S., Endo, S., & Ono, T. (2002). Generators of visual evoked potentials for faces and eyes in the human brain as determined by dipole localization. *Brain Topography, 15*(1), 51-63. doi: 10.1023/A:1019944607316

Taylor, M. J., Itier, R. J., Allison, T., & Edmonds, G. E. (2001). Direction of gaze effects on early face processing: eyes-only versus full faces. *Cognitive Brain Research*, *10*(3), 333–340. doi: 10.1016/S0926-6410(00)00051-3

Taylor, M.J., Edmonds, G.E., McCarthy, G., & Allison, T. (2001). Eyes first! Eye processing develops before face processing in children. *NeuroReport, 12*(8), 1671-1676. doi: 10.1097/00001756-200106130-00031

Vinette, C., Gosselin, F., & Schyns, P.G. (2004). Spatio-temporal dynamics of face recognition in a flash: It's in the eyes. *Cognitive Science, 28*(2), 289-301. doi: 10.1016/j.cogsci.2004.01.002

Willenbockel, V., Sadr, J., Fiset, D., Horne, G. O., Gosselin, F., & Tanaka, J. W. (2010). Controlling low-level image properties: the SHINE toolbox. *Behavior Research Methods*, *42*(3), 671–684. doi: 10.3758/BRM.42.3.671

Yarbus, A.L. (1967). *Eye movements and vision*. New York: Plenum Press.

Zheng, X., Mondloch, C.J., Nishimura, M., Vida, M.D., & Segalowitz, S.J. (2011). Telling one face from another: Electrocortical correlates of facial characteristics among individual female faces. *Neuropsychologia, 49*(12), 3254-3264. doi: 10.1016/j.neuropsychologia.2011.07.030