1

2 Revealing the neural time-course of direct gaze processing via spatial frequency manipulation

3 of faces

4 Inês Mares[a,b], Marie L. Smith[a,b], Mark H. Johnson[a] and Atsushi Senju[a]

5

6 [a]Centre for Brain and Cognitive Development, Department of Psychological Sciences,

7 Birkbeck, University of London

8 Henry Wellcome Building, Malet Street, London WC1E 7HX, United Kingdom

9

10 [b]Department of Psychological Sciences, Birkbeck, University of London

11 Birkbeck College, Malet Street, London WC1E 7HX, United Kingdom

12

13 Authors emails Inês Mares – imares01@mail.bbk.ac.uk; Marie L. Smith -

14 marie.smith@bbk.ac.uk; Mark H Johnson - mark.johnson@bbk.ac.uk and Atsushi Senju -

15 a.senju@bbk.ac.uk

16

17 Corresponding Author- Inês Mares,

18 Department of Psychological Sciences

19 Birkbeck, University of London

20 Malet St, London, WC1E 7HX

21 United Kingdom

1

2   Abstract

3   Direct gaze is a powerful social cue signalling the attention of another person toward oneself.

4   Here we investigated the relevance of low spatial frequency (LSF) and high spatial frequency

5   (HSF) in facial cues for direct gaze processing. We identified two distinct peaks in the ERP

6   response, the N170 and N240 components. These two components were related to different

7   stimulus conditions and influenced by different spatial frequencies. In particular, larger N170

8   and N240 amplitudes were observed for direct gaze than for averted gaze, but only in the

9   N240 component was this effect modulated by spatial frequency, where it was reliant in LSF

10   information. By contrast, larger N170 and N240 components were observed for faces than for

11   non-facial stimuli, but this effect was only modulated by spatial frequency in the N170

12   component, where it relied on HSF information. The present study highlights the existence of

13   two functionally distinct components related to direct gaze processing.

14   Keywords: Direct gaze; face processing; spatial frequency; event-related potentials; N170

15

16

17

18

19

20

21

1   Highlights:

2   -   Contribution of spatial frequency information to direct gaze processing was tested

3   -   Two early face-sensitive ERP components, N170 and N240, were identified

4   -   N170 reflected high spatial frequency processing of face vs non-face information

5   -   N170 enhancement to direct gaze motion is independent of spatial frequency content

6   -   N240 reflects low spatial frequency information for direct gaze processing

7

8

9

10

11

12

13

14

15

16

17

18

19

1. Introduction

Rapid and accurate perception of eye-gaze direction is an essential social function which permits the understanding of another's allocation of attention. Direct eye-gaze, in particular, signals when the attention of another is directed towards oneself and is therefore crucial for human social interaction throughout development to adulthood (Kleinke, 1986). In line with its importance, a number of visual search studies have suggested that direct gaze is a more salient visual cue than other gaze directions (Doi, Ueda, & Shinohara, 2009; Senju, Hasegawa, & Tojo, 2005; von Grünau & Anston, 1995), a phenomenon termed the stare-in-the-crowd effect. This faster orienting towards faces with direct gaze, compared with averted gaze, has further been shown in express saccades (saccades occurring under 130 ms), supporting the very early processing of gaze information (Mares, Smith, Johnson, & Senju, 2016). Accordingly, neural responses as observed through event related potentials (ERP) have shown that direct gaze is differentiated from averted gaze in early components such as the P100 (Berchio et al., 2016; Conty, Dezecache, Hugueville, & Grèzes, 2012) .

The importance of different spatial frequency bands for the fast processing of direct gaze remains, however, poorly understood. It has been proposed that coarser information, meaning low spatial frequencies (LSF), are important for fast object processing by activating faster magnocellular pathways reliant on this type of visual information (Bar, 2003, 2004). Early processing and detection of direct gaze in particular, has been suggested to rely on LSF information (Senju & Johnson, 2009). However, different studies have yielded very mixed results regarding the spatial frequencies that support gaze processing. The importance of LSF, for gaze processing has been shown in a gaze cueing paradigm (de Jong, van Engeland, & Kemner, 2008). de Jong et al. (2008) found a N200 ERP component effect, in neurotypical participants, for gaze cueing validity in which larger N200 amplitudes were observed for faces containing only LSF information compared with faces containing only more detailed

high spatial frequency (HSF) information. This pattern was reversed for participants with

autism spectrum disorders (ASD). Similarly, newborns, who lack sensitivity to HSFs (Norcia,

Tyler, & Hamer, 1990) and rely on LSF for face recognition (de Heering et al., 2008), are

sensitive to gaze information, looking longer towards faces with direct gaze than towards

faces with averted gaze (Farroni, Csibra, Simion, & Johnson, 2002).

On the other hand, it is not clear how the critical spatial frequency bandwidths for

gaze processing change throughout development, with HSF information being critical for the

explicit discrimination between leftwards and rightwards gaze directions in adults (Vida &

Maurer, 2015). Furthermore, early neural differentiation between direct and averted gaze has

only been observed for broad spatial frequency, while no effects of gaze direction were found

for stimulus displayed only with high or low spatial frequency information (Burra, Kerzel, &

George, 2016).

Thus, the present study aims to analyse the contribution of LSF and HSF for direct

gaze processing in an adapted version of a task that has previously shown to lead to enhanced

processing of direct gaze, both at neural and behavioural levels (Conty, N'Diaye, Tijus, &

George, 2007). The ERP paradigm used by Conty et al. (2007) has previously shown a larger

N170 component for dynamically presented direct gaze in comparison with averted gaze,

which was accompanied by a better recognition of direct gaze. Several studies have shown

the face selective N170 component, an early negativity occurring over lateral occipito-

temporal brain regions, is routinely and robustly associated with face perception and

encoding (e.g. Bentin, Allison, Puce, Perez, & McCarthy, 1996; Itier & Taylor, 2004) and is

sensitive to dynamic gaze motion (Conty et al., 2007; Latinus et al., 2015; Puce, Smith, &

Allison, 2000).

ERP studies analysing LSF and HSF modulation of the N170/M170 in face

processing, using different tasks, have found mixed results. While some studies have reported

no differences in this component between stimuli containing HSF or LSF information (Holmes, Winston, & Eimer, 2005; Pourtois, Dan, Grandjean, Sander, & Vuilleumier, 2005), other studies have reported either an advantage for stimuli containing only LSF information (Goffaux, Gauthier, & Rossion, 2003; Halit, de Haan, Schyns, & Johnson, 2006; Vlamings, Goffaux, & Kemner, 2009) or for stimuli containing only HSF information (Hsiao, Hsieh, Lin, & Chang, 2005; Nakashima et al., 2008). Task dependent effects have been observed to modulate N170 amplitudes to specific bandwidths of spatial frequency in faces (Goffaux, Jemel, Jacques, Rossion, & Schyns, 2003). Similarly, the use of different spatial frequency bands has been shown to be task dependent in a previous behavioural study (Schyns & Oliva, 1999). Thus, to try to avoid a direct impact of task, the present study only asked participants to detect a rarely occurring target to keep participants attentive.

Original images of faces with direct and averted gaze were filtered to create stimuli presenting high, low and broad (unfiltered) spatial frequencies. Furthermore, matched conditions were created with pictures of buildings, which allowed for the dissociation between general motion processing and eye gaze processing. As a secondary question, the current study also allowed us to analyse separately the impact of HSF and LSF information on the ERP components associated with facial motion compared to non-facial motion, and those associated with different direction of gaze shift.

2. Methods

2.1. Participants

Fifteen participants (10 female, between 21-48 years old, mean age: 33.13 years) were tested. They were right handed, with normal or corrected to normal vision. All participants signed a written informed consent form and were paid a small sum or received course credits

for their participation. The study was approved by the ethical committee of the Department of Psychological Sciences, Birkbeck College, University of London.

2.2. Stimuli

Twenty-four greyscale digitized photographs were used as stimuli, 12 of which were of faces (6 male and 6 female) and the other 12 were of buildings. Faces had neutral expressions and were cropped to exclude external features such as hair. A similar oval cropping was used for the building stimuli. Faces with direct and averted (30° of center, counterbalanced between right and left) gaze directions were included (see George et al., 2001 for details about the stimuli generation process). Faces were presented oriented to the left and to the right (30°), to avoid a symmetry confound. To be able to simulate a change in gaze direction, a morphed version of direct and averted gaze was created to establish a common baseline for both conditions (eye-gaze at approximately 15°). Similarly, a black bar was inserted in building images either centred, deviated to the left or to the right or morphed between them (Figure 1a). This control was used to account for differences in the perception of motion to the centre or towards the periphery. HSF (Figure 1b) and LSF (Figure 1c) versions of all images were created by applying a high (above 24 cycles per stimuli) and low (below 8 cycles per stimuli) pass filter. Second order Butterworth filters were used to filter stimuli, using an in house Matlab script (The Mathworks, Natick, MA). The selected spatial frequency cut-offs per image and filtering were similar to the ones used by Schyns and Oliva (1999). These spatial frequency cut-offs are outside the mid-range spatial frequencies considered crucial for face recognitions (8-16 Hz) and have been previously used in other ERP studies tackling the spatial frequency information used for face processing (Halit, de Haan, & Johnson, 2003). Images without any filtering were also presented (BSF). Mean luminance and root mean square contrast (RMS) were equalized between all conditions. RMS

1 has been shown to be the best index for perceived contrast in natural images (Bex & Makous,

2 2002). Stimuli subtended 14°×18° degrees of visual angle.

3 *2.3. Procedure*

4 Participants sat comfortably in a chair in an electrically shielded and soundproofed

5 room at a fixed viewing distance of 60 cm from the computer screen (fixed via the use of a

6 chin rest). Stimuli were presented with E-Prime software (Psychology Software Tools,

7 Pittsburgh, PA, USA). Trials began with a fixation cross presented for 500 ms at the

8 horizontal midline and at the height equivalent to the eye region of the face stimuli. Illusion

9 of movement was accomplished by the successive presentation of a morphed stimulus,

10 followed by the same stimulus in its direct or averted version. The initial intermediate gaze

11 stimuli appeared for a random duration between 600-800 ms followed by its version in direct

12 or averted gaze for 500 ms. The same sequence was used with buildings as a non-social

13 stimulus to control for the direction of perceived motion. Overall participants performed 6

14 blocks of 156 trials. Each stimulus sequence was randomly presented in its direct/centred and

15 averted/deviated conditions, and in the three spatial frequency conditions: broad, high and

16 low. An equal number of building and face trials were shown counterbalanced between

17 direct/centred and averted/deviated conditions. To maintain attention participants were

18 instructed to press a key upon the appearance of a red square superimposed on the stimuli.

19 This occurred in 11.1% of the trials, which were excluded from further analysis. Note that red

20 squares were not filtered, and thus were always presented in broad spatial frequency.

21 *2.4. ERP recording and Data analyses*

22 Electroencephalographic (EEG) activity was continuously recorded from 60 Ag/AgCl

23 electrodes placed according to the international 10/10 system on a fitted cap (EASYCAP,

24 Herrsching-Breitbrunn, Germany). Two EOG electrodes were placed on the outer canthi of

the eyes to detect horizontal eye movements and a third was placed underneath one eye to detect vertical eye movements and blinks. Data was online referenced to the mastoids and acquired at a sampling rate of 1000 Hz. Electrode impedance was kept below 10 kΩ. Data analysis was performed via the Matlab (The Mathworks, Inc., Natick, MA, USA) toolbox EEGLAB (Delorme & Makeig, 2004). Data was offline band pass filtered between 0.1 and 40 Hz and an average reference was used. Epochs were created around the second image onset (-100 to 600 ms). Large artefacts or saccades were removed trough visual inspection prior to the use of Independent Components Analysis (ICA) to remove blinks. Number of removed components varied between participants with zero components removed in 5 participants due to lack of blinks, 1 component removed in 9 participants, and 2 components removed in 1 participant to completely remove blinks. An average of 18.6% trials were rejected per participant. Epochs were baseline corrected using the 100 ms prior to the onset of the second face/building.

Two early negative peaks were identified in occipital-temporal areas bilaterally. Channels in occipital-temporal areas were chosen where these peaks were maximal in amplitude (TP7/8, P7/8, PO7/8, Figure 2). The existence of two peaks around this time window has previously been described in a study of biological motion as the N170 and the N240 (Hirai, Fukushima, & Hiraki, 2003). To strengthen the observation and identification of the two distinct components, we used the global field power metric (GFP, Skrandies, 1990) to identify them and select the appropriate time windows (Figure 3; Supplementary material). GFP computes the voltage spatial standard deviation from each electrode's voltage at each time sample. This provides a single measure for each time point in which a larger signal standard deviation corresponds to more signal strength. The same two separate components were identified via the GFP and time windows were created to encompass them around 50% of each components' amplitude - from 155 to 189 ms for N170 and 225 to 270 ms for N240.

Due to the type of paradigm, in which epochs are centred around the second image onset, the P100 component was very diminished. Nonetheless, an early time window from 90 to 125 ms was also considered to analyse possible early effects of gaze direction in the P100 component. This analysis was performed in the previously mentioned occipital-temporal electrodes (TP7/8, P7/8, PO7/8).

Mean amplitudes for all time windows were calculated per participant, from the pooled data of the selected channels in each hemisphere (Figure 4 for the ERP waveforms). Statistical analysis of the mean amplitudes was carried out using a four-way repeated measures ANOVA with hemisphere (right and left), spatial frequency (broad, low, high), category of stimulus (face and buildings) and direction of motion (towards the centre or periphery) as within-subject factors. When appropriate, post hoc planned comparisons were performed. Violations of sphericity were corrected with the Greenhouse-Geisser correction. Effect sizes for dependent t-tests were calculated using the formula proposed in Eq(3) of Dunlap, Cortina, Vaslow & Burke (1996), to correct for an overestimation of the effect size for paired comparisons.

3. Results

3.1. Behavioural results

Accuracy on the detection task was at ceiling level (M=94±3%), with no effect of spatial frequency (F[2,28]=0.62, p=.544, η_p^2=.043), category of stimulus (F[1,14]=0.02, p=.902, η_p^2=.001), or direction of motion (F[1,14]=0.02, p=.890, η_p^2=.001), or an interaction between any of them.

3.2. ERP results

3.2.1. P100

An interaction between category (face vs buildings) and direction of motion (centre vs periphery) was not observed ($F[1,14]=.40$, $p=.539$, $\eta_p^2=.03$), nor between these two factors and special frequency ($F[1,14]=1.56$, $p=.229$, $\eta_p^2=.100$), showing that no effect of gaze was found in this component. There was only an effect of stimulus category ($F[1,14]=7.48$, $p=.016$, $\eta_p^2=.348$), driven by larger P100 for faces (M=.05 µV) than buildings (M=-.16 µV).

3.2.2. N170

For the N170 component, main effects of hemisphere ($F[1,14]=14.50$, $p=.002$, $\eta_p^2=.509$) and spatial frequency ($F[2,28]=15.34$, $p<.001$, $\eta_p^2=.523$) were significant, driven by larger responses over the right hemisphere (M=-2.00) than the left (M=-.87; p=.002) and smaller responses for LSF stimuli (M=-1.08µV) compared with HSF (M=-1.51µV; p=.004) and BSF (M=-1.72µV; p<.001).

Furthermore, an interaction between the category (face vs buildings) and direction of motion (centre vs periphery) was observed ($F[1,14]=6.25$, $p=.025$, $\eta_p^2=.309$), which did not interact with spatial frequency content ($F[2,28]=1.05$, $p=.364$, $\eta_p^2=.070$). Faces with direct gaze (M=-1.66µV) showed a significantly larger negative amplitude than faces with averted gaze (Figure 5, M=-1.41 ±.85 µV; $t[14]=2.47$, $p=.027$, *Cohen's d*=.274). No effect was found between the direction of motion in the two building conditions ($t[14]=-1.52$, $p=.151$, *Cohen's d*= -.193). Differences between faces with direct gaze and buildings with centred bar were marginal ($t[14]=2.10$, $p=.054$, *Cohen's d*=.444). No significant difference was observed between averted gaze and buildings with deviated bar ($t[14]=-.064$, $p=.950$, *Cohen's d*=-.012). An interaction between spatial frequency and category modulated these effects, ($F[2,28]=6.63$, $p=.004$, $\eta_p^2=.321$), with an enhanced processing of high spatial frequency face stimuli (vs. buildings, Figure 6, M= -1.81µV vs. M= -1.22 µV, $t[14]=-3.98$, $p=.001$, *Cohen's*

1   *d*=-.632), in the absence of category effects for BSF(*t*[14]=-.53, *p*=.602, *Cohen's d*=-.149)

2   and LSF stimuli (*t*[14]=1.09, *p*=.296, *Cohen's d*=.188).

3   *3.2.3. N240*

4   Main effects of hemisphere (*F*[1,14]=4.93, *p*=043., η_p^2=.261), spatial frequency

5   (*F*[2,28]=7.88, *p*=.002, η_p^2=.360) and category were observed (*F*[1,14]=30.644, *p*<.001,

6   η_p^2=.686) for the N240 component, where larger amplitudes were observed for the right

7   hemisphere (M=-1.66µV) than the left (M=-1.13µv; p=.043), for BSF stimuli (M=-1.75 µV)

8   than HSF (M=-1.06µV; p=.001) and LSF (M=-1.38 µv; p=.047) and for faces (M=-1.93 µv)

9   than buildings (M=-.86 µv; *p*<.001). Importantly, a three way interaction between spatial

10   frequency, category of stimulus and direction of motion (Figure7, *F*[2,28]=3.964, *p*=.030,

11   η_p^2=.221) was observed. To interpret this interaction, separate two-way ANOVAS were

12   conducted to analyse the interaction between category and direction of motion in each spatial

13   frequency band.

14   In BSF a main effect of category was observed (*F*[1,14]=13.47, *p*=.003, η_p^2=.490)

15   with a larger negativity for faces (M= -2.31µV) than buildings (M= -1.19µV). No effect of

16   direction of motion (*F*[1,14]=3.93, *p*=.067, η_p^2=.219), or an interaction between these two

17   factors was found (*F*[1,14]=.10, *p*=.921, η_p^2=.001).

18   Similarly in HSF a main effect of category was observed (*F*[1,14]=64.71, *p*<.001, η_p

19   ^{2}=.822) with a larger negativity for faces (M= -1.75µV) than buildings (M= -.37µV), in the

20   absence of an effect of direction of motion (*F*[1,14]=0.75, *p*=.402, η_p ^{2}=.051), or an

21   interaction between them (*F*[1,14]=0.001, *p*=.973, η_p^2<.001).

22   In LSF, we observed a main effect of category (*F*[1,14]=7.56, *p*=.016, η_p^2=.351; with

23   a larger negativity for faces [M=-1.73µv] than buildings [M= -1.04µV]), and no main effect

of direction motion (F[1,14]=3.93, p=.067, η_p^2=.219). However, we also observed a two-way interaction between these factors (F [1,14]=10.08, p=.007, η_p^2=.418). Crucially, faces moving to direct gaze (M= -1.98µV) elicited larger N240 amplitudes than faces moving to averted gaze (M=-1.47µV; t[14]=2.44, p=.029, *Cohen's d*=.386). Unexpectedly, a difference between the direction of motion in the two building stimuli was also observed (t[14]=-2.47, p=.027, *Cohen's d*=-.411). Importantly this difference was in the opposite direction, with buildings with a bar moving towards the periphery showing a larger negative amplitude (M=-1.23 µV) than buildings with a bar moving towards the centre (M=-.83 µV). A comparison between faces and buildings with motion towards the centre showed a larger N240 for faces with direct gaze than for buildings with a centred bar *(t*[14]=3.54, p=.003, *Cohen's d*=1.23), while no difference was found between faces and buildings with motion towards the periphery (t[14]=.921, p=.373, *Cohen's d*=.335).

4. Discussion

The present study analysed the importance of LSF and HSF information for the differential processing of direct gaze. Our findings suggest that gaze is processed in two stages, an initial step, corresponding to the N170 component, that is not specifically modulated by the spatial frequency content of the stimuli and a later step, corresponding to the N240 component, which is dependent on LSF information. By contrast, larger N170 amplitudes for faces compared to non-faces relied on HSF information, while the larger N240 amplitudes for faces was not modulated by spatial frequency manipulation.

To our knowledge, the presence of a bifid N2 (i.e. composed by two distinct subcomponents) has not been previously described in studies analysing gaze modulation of the N170 component in adults. Despite this, several studies have shown the time window similar to N240, which immediately follows N170, to be relevant for gaze processing (Conty

et al., 2007; Myllyneva & Hietanen, 2015; Pönkänen, Alhoniemi, Leppänen, & Hietanen, 2011). Furthermore, although not directly reported as such, a similar bifid structure over occipito-temporal areas can also be seen in the results of Conty's (2007) study, were a larger amplitude for faces with direct gaze was observed over occipito-temporal electrodes between 160 and 210 ms. Other ERP studies that analysed gaze processing have looked at occipito-temporal electrodes activity in an early time window following N170 (e.g. 160-300 ms Myllyneva & Hietanen, 2015; 190-290 ms, Pönkänen, Alhoniemi, Leppänen, & Hietanen, 2010), and found differences between gaze direction in these periods. Taken together these findings support the existence of a second component critical for gaze processing, especially when considering the bifid nature of the GFP wave presently found.

4.1. N170 component

Overall, larger amplitudes were observed in the N170 component for HSF and BSF than LSF stimuli, possibly related with the importance of HSF for face specific processing. This is consistent with studies reporting larger N170 amplitudes for stimuli containing only HSF information (Hsiao et al., 2005; Nakashima et al., 2008). Nonetheless several studies have reported opposite findings with larger N170/M170 for LSF than HSF stimuli (Goffaux et al., 2003; Goffaux, Jemel, Jacques, Rossion, & Schyns, 2003; Halit, de Haan, Schyns, & Johnson, 2006). Face specific N170 effects in particular, have yielded mixed results either supporting their reliance on LSF information (Goffaux, Gauthier, et al., 2003), HSF information (Nakashima et al., 2008), or reporting no difference between spatial frequencies (Holmes et al., 2005). The use of different spatial frequency bands has been shown to be task dependent (Schyns & Oliva, 1999) and as observed with the N170 component (Goffaux, Jemel, et al., 2003). Similarly, in the present findings, while HSF information seems critical for face specific effects in the N170, this was not the case for gaze specific effects.

A limitation of the present study is the use of physically different stimuli, which does not allow to completely disregard possible low level confounds, even though we tried to minimise this issue by matching both contrast and luminance across all categories. Note that no effect of gaze or spatial frequency were found in the P100, the component which would be sensitive to low level features, suggesting that low level confound, if any, would be negligible.

4.2. Gaze effects in N170 and N240

A larger N170 negativity for faces with direct gaze than faces with averted gaze replicates the original findings of Conty et al. (2007). Larger amplitudes for direct gaze in this component have been associated with a larger Superior Temporal Sulcus (STS) activation for direct gaze (Conty et al., 2007). These results are also in line with other imaging studies using dynamic gaze shifts that showed a larger STS activation for direct gaze (Ethofer, Gschwind, & Vuilleumier, 2011; Kuzmanovic et al., 2009; Pelphrey, Viola, & McCarthy, 2004). Although it can be argued that the differences found for gaze direction could be linked to the incongruence between direct gaze and averted head position, this seems unlikely given the similar pattern of results for both frontal and oriented faces in an earlier study with a comparable paradigm (Conty et al., 2007). Furthermore, the larger amplitudes for direct gaze shown in the current study cannot be fully attributed to low-level visual features such as direction of motion, since the observed effects in the buildings' conditions occurred in the opposite direction. The larger amplitude in the N240 component observed for LSF buildings motion towards the periphery compared with the centre was not predicted. This might suggest that the N240 component is particularly tuned to motion perception, including from non-biological sources. Differences between gaze direction in

N200 have been previously shown to be mediated by LSF but not HSF (de Jong et al., 2008). In their experiment, faces initially in direct gaze dynamically changed their gaze leftwards or rightwards, cueing a target in 50% of the trials. In typical participants, they found a larger N200 to LSF faces than HSF faces to correctly cued targets, although this pattern was reversed in participants with autism spectrum disorder (ASD). Combined with the current results, LSF information could be relevant for implicit processing of direct gaze, as well as gaze cueing (de Jong et al., 2008) in a second stage of processing.

Interestingly, for LSF stimuli in the N240 peak, only direct gaze faces, and not averted gaze faces, differed from their building stimuli counterparts. Similarly, larger N170 amplitudes for faces compared with buildings were marginally observed for faces with direct gaze but were absent for faces with averted gaze. These results give some support to the proposal that faces with direct gaze are a better match to a face template due to the contrasts between iris and sclera in the eye region (Senju & Johnson, 2009).

4.3. *Two stages of direct gaze processing*

The current results support the existence of at least two stages of gaze processing sensitive to different spatial frequencies. An initial component, N170 was sensitive to direct gaze information but not dependent on low spatial frequency information, while differentiation between gaze directions in a second component, N240, seemed to rely on LSF. This could signify that an initial stage of gaze processing observable by surface EEG relied mostly on cortical pathways independently of magnocellular input, while a later stage, occurring around 225 ms relied on LSF information which could possibly be mediated by amygdala activation (Morris, DeGelder, Weiskrantz, & Dolan, 2001; Vuilleumier, Armony, Driver, & Dolan, 2003). In support of the role of the amygdala for this second component, amygdala activation to gaze or specific to the eye region has been observed to occur around

200 ms after stimulus onset (Dumas et al., 2013; Meletti et al., 2012; Sato et al., 2011 but see Huijgen et al., 2015 for a longer cluster of amygdala activation starting from 123 ms until 258 ms). Furthermore, considering both the effects on face and gaze processing, these findings are consistent with those reported by Vuilleumier et al. (2003). In Vuilleumier et al (2003) fMRI study, a face identity repetition effect was found in the fusiform gyrus (an area linked with face processing and the N170 component) mediated by HSF while an emotion effect was observed in the amygdala driven by LSF information.

Nonetheless, the present data does not provide support for the hypothesis that direct gaze information transmitted through magnocellular inputs enhances face perception, which would be associated with a fast modulation of earlier components of the ERP. Due to the nature of the present paradigm, in which the ERPs are associated with the appearance of the gaze shift, there was a very small P100 component. No difference between gaze direction was identified in the P100, as has been observed in other studies (Berchio et al., 2016; Conty, Dezecache, Hugueville, & Grèzes, 2012).

Nonetheless, the current results showed that the selective gaze effect on the LSF condition was observed later, in the N240 component. To understand the neural underpinnings of this later effect further imaging and/or lesion studies are needed. It will also be crucial to develop experimental paradigms that are reliably sensitive to gaze direction in earlier components than the N170, to test whether these earlier ERP components could be modulated by the LSF information.

5. Conclusion

The present study suggests that processing of direct gaze occurs in two stages with the later one being modulated by LSF information. Overall, HSF information appears to be particularly important for discriminating between face and non-face objects over the N170

component, while LSF information appears to be crucial for gaze discrimination over the N240 component. Future imaging studies combining EEG and fMRI should be conducted to analyse the underlying structures mediating these two stages of gaze processing as well as their function.

6. Acknowledgements

7. Conflicts of interest

Authors have no conflicts of interest.

8. Bibliography

Bar, M. (2003). A cortical mechanism for triggering top-down facilitation in visual object recognition. *Journal of Cognitive Neuroscience*, *15*(4), 600–609. https://doi.org/10.1162/089892903321662976

Bar, M. (2004). Visual objects in context. *Nature Reviews. Neuroscience*, *5*(8), 617–29. https://doi.org/10.1038/nrn1476

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. https://doi.org/10.1162/jocn.1996.8.6.551

Berchio, C., Rihs, T. A., Piguet, C., Dayer, A. G., Aubry, J.-M., & Michel, C. M. (2016). Early averted gaze processing in the right Fusiform Gyrus: An EEG source imaging study. *Biological Psychology*. https://doi.org/10.1016/j.biopsycho.2016.06.008

Bex, P. J., & Makous, W. (2002). Spatial frequency, phase, and the contrast of natural

images. *Journal of the Optical Society of America. A, Optics, Image Science, and Vision*, *19*(6), 1096–1106. https://doi.org/10.1364/JOSAA.19.001096

Burra, N., Kerzel, D., & George, N. (2016). Early left parietal activity elicited by direct gaze: A high-density EEG study. *PLoS ONE*, *11*(11), 1–23. https://doi.org/10.1371/journal.pone.0166430

Conty, L., Dezecache, G., Hugueville, L., & Grèzes, J. (2012). Early binding of gaze, gesture, and emotion: neural time course and correlates. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *32*(13), 4531–9. https://doi.org/10.1523/JNEUROSCI.5636-11.2012

Conty, L., N'Diaye, K., Tijus, C., & George, N. (2007). When eye creates the contact! ERP evidence for early dissociation between direct and averted gaze motion processing. *Neuropsychologia*, *45*(13), 3024–3037. https://doi.org/10.1016/j.neuropsychologia.2007.05.017

de Heering, A., Turati, C., Rossion, B., Bulf, H., Goffaux, V., & Simion, F. (2008). Newborns' face recognition is based on spatial frequencies below 0.5 cycles per degree. *Cognition*, *106*(1), 444–54. https://doi.org/10.1016/j.cognition.2006.12.012

de Jong, M. C., van Engeland, H., & Kemner, C. (2008). Attentional effects of gaze shifts are influenced by emotion and spatial frequency, but not in autism. *Journal of the American Academy of Child and Adolescent Psychiatry*, *47*(4), 443–454. https://doi.org/10.1097/CHI.0b013e31816429a6

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. https://doi.org/10.1016/j.jneumeth.2003.10.009

Doi, H., Ueda, K., & Shinohara, K. (2009). Neural correlates of the stare-in-the-crowd effect. *Neuropsychologia*, *47*(4), 1053–1060. https://doi.org/10.1016/j.neuropsychologia.2008.11.004

Dumas, T., Dubal, S., Attal, Y., Chupin, M., Jouvent, R., Morel, S., & George, N. (2013). MEG Evidence for Dynamic Amygdala Modulations by Gaze and Facial Emotions. *PLoS ONE*, *8*(9), 1–11. https://doi.org/10.1371/journal.pone.0074145

Dunlap, W. P., Cortina, J. M., Vaslow, J. B., & Burke, M. J. (1996). Meta-analysis of experiments with matched groups or repeated measures designs. *Psychological Methods*, *1*(2), 170–177. https://doi.org/10.1037/1082-989X.1.2.170

Ethofer, T., Gschwind, M., & Vuilleumier, P. (2011). Processing social aspects of human gaze: a combined fMRI-DTI study. *NeuroImage*, *55*(1), 411–19. https://doi.org/10.1016/j.neuroimage.2010.11.033

Farroni, T., Csibra, G., Simion, F., & Johnson, M. H. (2002). Eye contact detection in humans from birth. *Proceedings of the National Academy of Sciences of the United States of America*, *99*(14), 9602–9605. https://doi.org/10.1073/pnas.152159999

George, N., Driver, J., & Dolan, R. J. (2001). Seen gaze-direction modulates fusiform activity and its coupling with other brain areas during face processing. *NeuroImage*, *13*(6 Pt 1), 1102–12. https://doi.org/10.1006/nimg.2001.0769

Goffaux, V., Gauthier, I., & Rossion, B. (2003). Spatial scale contribution to early visual differences between face and object processing. *Cognitive Brain Research*, *16*(3), 416–

424. https://doi.org/10.1016/S0926-6410(03)00056-9

Goffaux, V., Jemel, B., Jacques, C., Rossion, B., & Schyns, P. G. (2003). ERP evidence for task modulations on face perceptual processing at different spatial scales. *Cognitive Science*, *27*(2), 313–325. https://doi.org/10.1016/S0364-0213(03)00002-8

Halit, H., de Haan, M., & Johnson, M. . H. (2003). Cortical specialisation for face processing: face-sensitive event-related potential components in 3- and 12-month-old infants. *NeuroImage*, *19*(3), 1180–1193. https://doi.org/10.1016/S1053-8119(03)00076-4

Halit, H., de Haan, M., Schyns, P. G., & Johnson, M. H. (2006). Is high-spatial frequency information used in the early stages of face detection? *Brain Research*, *1117*(1), 154–161. https://doi.org/10.1016/j.brainres.2006.07.059

Hirai, M., Fukushima, H., & Hiraki, K. (2003). An event-related potentials study of biological motion perception in humans. *Neuroscience Letters*, *344*(1), 41–44. https://doi.org/10.1016/S0304-3940(03)00413-0

Holmes, A., Winston, J. S., & Eimer, M. (2005). The role of spatial frequency information for ERP components sensitive to faces and emotional facial expression. *Brain Research. Cognitive Brain Research*, *25*(2), 508–20. https://doi.org/10.1016/j.cogbrainres.2005.08.003

Hsiao, F. J., Hsieh, J. C., Lin, Y. Y., & Chang, Y. (2005). The effects of face spatial frequencies on cortical processing revealed by magnetoencephalography. *Neuroscience Letters*, *380*(1–2), 54–59. https://doi.org/10.1016/j.neulet.2005.01.016

Huijgen, J., Dinkelacker, V., Lachat, F., Yahia-Cherif, L., El Karoui, I., Lemaréchal, J.-D., … George, N. (2015). Amygdala processing of social cues from faces: an intracrebral EEG study. *Social Cognitive and Affective Neuroscience*, nsv048-. https://doi.org/10.1093/scan/nsv048

Itier, R. J., & Taylor, M. J. (2004). Source analysis of the N170 to faces and objects. *Neuroreport*, *15*(8), 1261–1265. https://doi.org/10.1097/01.wnr.0000127827.73576.d8

Kleinke, C. L. (1986). Gaze and eye contact: a research review. *Psychological Bulletin*, *100*(1), 78–100. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/3526377

Kuzmanovic, B., Georgescu, A. L., Eickhoff, S. B., Shah, N. J., Bente, G., Fink, G. R., & Vogeley, K. (2009). Duration matters: dissociating neural correlates of detection and evaluation of social gaze. *NeuroImage*, *46*(4), 1154–1163. https://doi.org/10.1016/j.neuroimage.2009.03.037

Latinus, M., Love, S. A., Rossi, A., Parada, F. J., Huang, L., Conty, L., … Puce, A. (2015). Social decisions affect neural activity to perceived dynamic gaze. *Social Cognitive and Affective Neuroscience*, *10*(11), 1557–1567. https://doi.org/10.1093/scan/nsv049

Mares, I., Smith, M. L., Johnson, M. H., & Senju, A. (2016). Direct gaze facilitates rapid orienting to faces: Evidence from express saccades and saccadic potentials. *Biological Psychology*, *121*, 84–90. https://doi.org/10.1016/j.biopsycho.2016.10.003

Meletti, S., Cantalupo, G., Benuzzi, F., Mai, R., Tassi, L., Gasparini, E., … Nichelli, P. (2012). Fear and happiness in the eyes: An intra-cerebral event-related potential study from the human amygdala. *Neuropsychologia*, *50*(1), 44–54. https://doi.org/10.1016/j.neuropsychologia.2011.10.020

Morris, J. S., DeGelder, B., Weiskrantz, L., & Dolan, R. J. (2001). Differential

extrageniculostriate and amygdala responses to presentation of emotional faces in a cortically blind field. *Brain : A Journal of Neurology*, *124*(Pt 6), 1241–52. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11353739

Myllyneva, A., & Hietanen, J. K. (2015). There is more to eye contact than meets the eye. *Cognition*, *134*, 100–109. https://doi.org/10.1016/j.cognition.2014.09.011

Nakashima, T., Kaneko, K., Goto, Y., Abe, T., Mitsudo, T., Ogata, K., … Tobimatsu, S. (2008). Early ERP components differentially extract facial features: Evidence for spatial frequency-and-contrast detectors. *Neuroscience Research*, *62*(4), 225–235. https://doi.org/10.1016/j.neures.2008.08.009

Norcia, A., Tyler, C., & Hamer, R. (1990). Development of Contrast Sensitivity in the Human Infant. *Vision Research*, *30*(10), 1475–1486. https://doi.org/10.1016/0042-6989(90)90028-J

Pelphrey, K. A., Viola, R. J., & McCarthy, G. (2004). When strangers pass: processing of mutual and averted social gaze in the superior temporal sulcus. *Psychological Science*, *15*(9), 598–603. https://doi.org/10.1111/j.0956-7976.2004.00726.x

Pönkänen, L. M., Alhoniemi, A., Leppänen, J. M., & Hietanen, J. K. (2011). Does it make a difference if I have an eye contact with you or with your picture? An ERP study. *Social Cognitive and Affective Neuroscience*, *6*(4), 486–494. https://doi.org/10.1093/scan/nsq068

Pourtois, G., Dan, E. S., Grandjean, D., Sander, D., & Vuilleumier, P. (2005). Enhanced extrastriate visual response to bandpass spatial frequency filtered fearful faces: Time course and topographic evoked-potentials mapping. *Human Brain Mapping*, *26*(1), 65–79. https://doi.org/10.1002/hbm.20130

Puce, A., Smith, A., & Allison, T. (2000). Erps evoked by viewing facial movements. *Cognitive Neuropsychology*, *17*(1), 221–39. https://doi.org/10.1080/026432900380580

Sato, W., Kochiyama, T., Uono, S., Matsuda, K., Usui, K., Inoue, Y., & Toichi, M. (2011). Rapid Amygdala Gamma Oscillations in Response to Eye Gaze. *PLoS ONE*, *6*(11), e28188. https://doi.org/10.1371/journal.pone.0028188

Schyns, P. G., & Oliva, A. (1999). Dr. Angry and Mr. Smile: when categorization flexibly modifies the perception of faces in rapid visual presentations. *Cognition*, *69*(3), 243–265. https://doi.org/10.1016/S0010-0277(98)00069-9

Senju, A., Hasegawa, T., & Tojo, Y. (2005). Does perceived direct gaze boost detection in adults and children with and without autism? The stare-in-the-crowd effect revisited. *Visual Cognition*, *12*(8), 1474–1496. https://doi.org/10.1080/13506280444000797

Senju, A., & Johnson, M. H. (2009). The eye contact effect: mechanisms and development. *Trends in Cognitive Sciences*, *13*(3), 127–134. https://doi.org/10.1016/j.tics.2008.11.009

Skrandies, W. (1990). Global field power and topographic similarity. *Brain Topography*, *3*(1), 137–141. https://doi.org/10.1007/BF01128870

Vida, M. D., & Maurer, D. (2015). A comparison of spatial frequency tuning for judgments of eye gaze and facial identity. *Vision Research*, *112*, 45–54. https://doi.org/10.1016/j.visres.2015.04.018

Vlamings, P. H. J. M., Goffaux, V., & Kemner, C. (2009). Is the early modulation of brain activity by fearful facial expressions primarily mediated by coarse low spatial frequency

information? *Journal of Vision*, *9*(5), 12.1-13. https://doi.org/10.1167/9.5.12

von Grünau, M., & Anston, C. (1995). The detection of gaze direction: a stare-in-the-crowd effect. *Perception*, *24*(11), 1297–1313. https://doi.org/10.1068/p241297

Vuilleumier, P., Armony, J. L., Driver, J., & Dolan, R. J. (2003). Distinct spatial frequency sensitivities for processing faces and emotional expressions. *Nature Neuroscience*, *6*(6), 624–631. https://doi.org/10.1038/nn1057
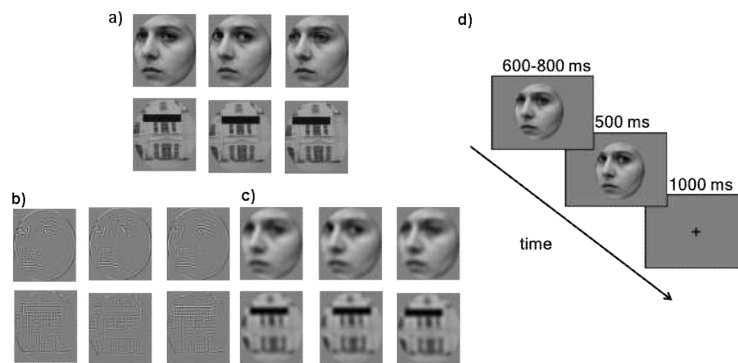
Figure 1 – Examples of face and building stimuli in a) BSF, b) HSF and c) LSF. d) Schematics of the procedure.

1

2

3

4

5

6

BSF

HSF

LSF

Buildings Away
Buildings Central
Averted Gaze
Direct Gaze

TP7 P7 P5 P3 PO7 PO3 POz PO4 PO8 P4 P6 P8 TP8 O1 Oz O2

1 μv
-100   500
-4
Time (ms)
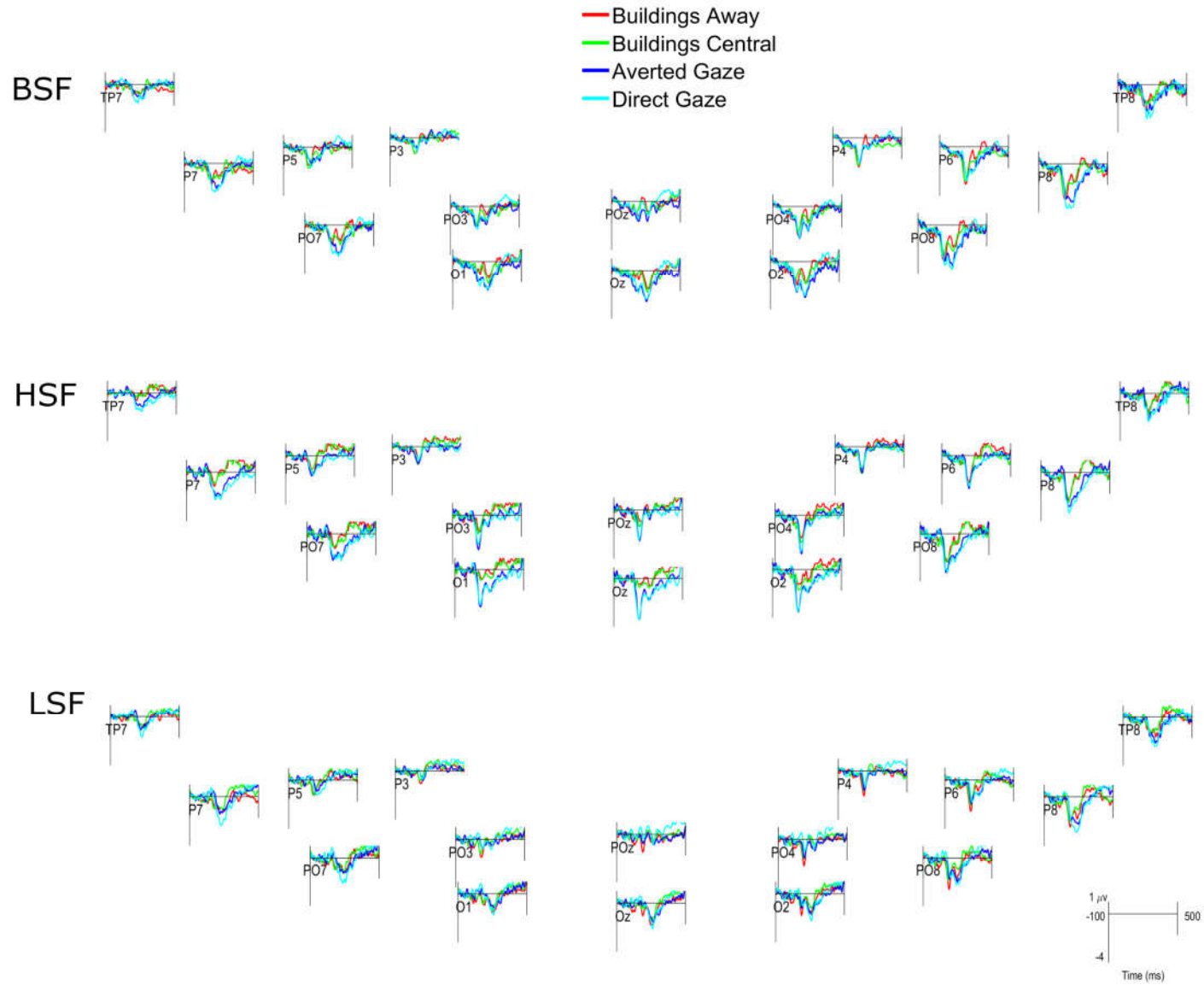
1   Figure 2 – Grand average ERP waveforms for each condition, with direct and averted gaze as

2   well as buildings with movement towards the centre or periphery, across spatial frequencies.

3   All electrodes over occipito-temporal areas are shown (TP7/8, P7/8, P5/6, PO7/8, P3/4,

4   PO3/4, O1/O2, POz and Oz).

5

6

7

8

9

10
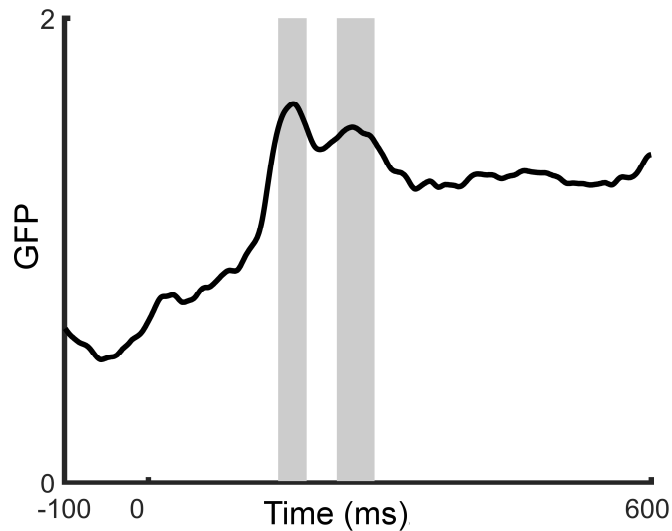
11

12

13   Figure 3- Timecourse of global field power averaged for all conditions over all electrodes.

14   The two grey time windows correspond to N170 and N240.
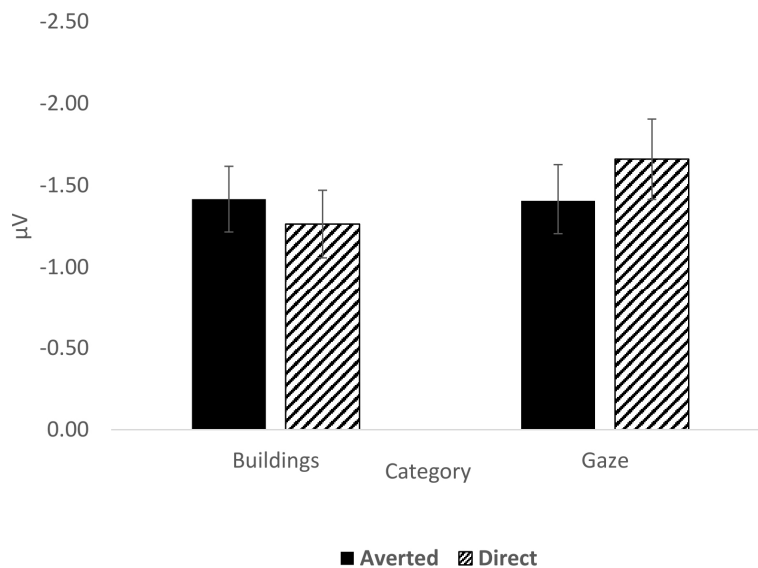
15

16

17

18

1

2

3

4
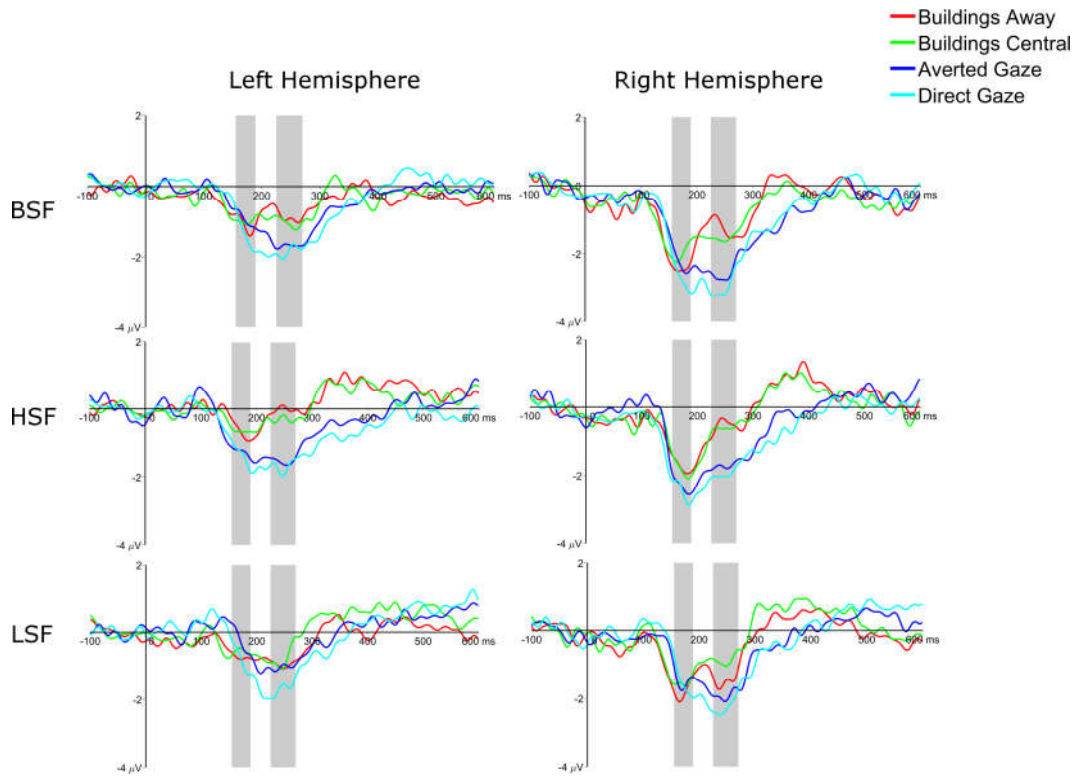
5 Figure 4- Grand averaged ERPs for faces with direct and averted gaze as well as buildings

6 with movement towards the centre or periphery. Grand averages were taken from 6 electrodes

7 (TP7/8, P7/8, PO7/8) for each spatial frequency considered. Shadowed areas corresponded to
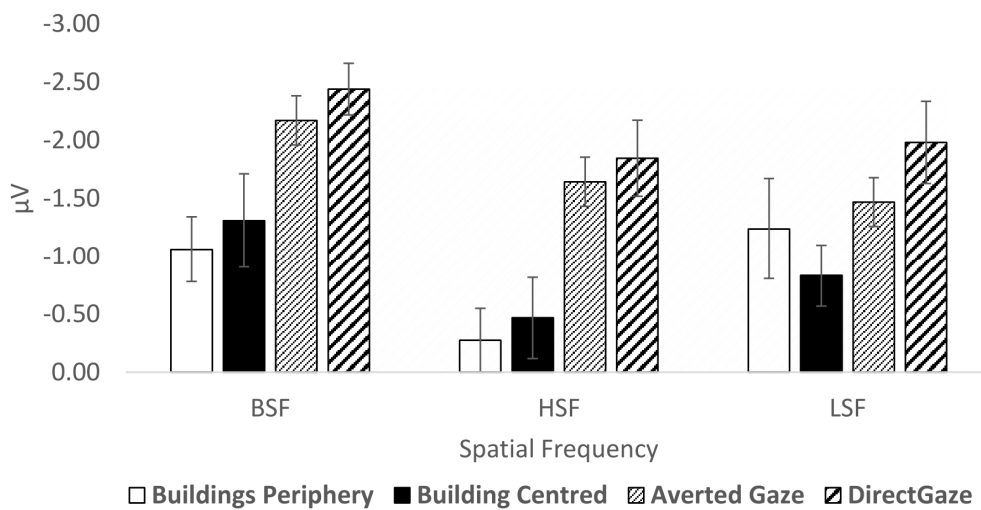
8 the time windows analysed for N170 and N240.

9

10 Figure 5 - Mean amplitudes of the N170 across categories from electrodes TP7/8, P7/8, P5/6,

11 PO7/8, P3/4, PO3/4, O1/O2, POz and Oz. Error bars with standard error.
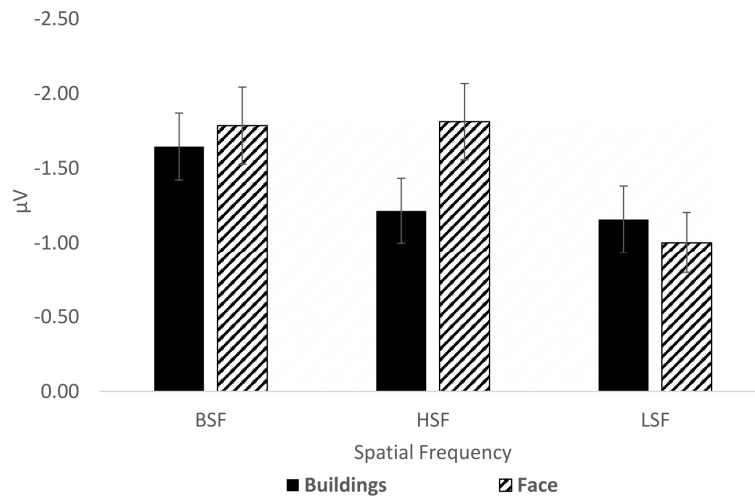
Figure 6 - Mean amplitudes of the N170 across spatial frequencies from electrodes TP7/8, P7/8, P5/6, PO7/8, P3/4, PO3/4, O1/O2, POz and Oz. Error bars with standard error.

Figure 7 - Mean amplitudes of the N240 across categories and spatial frequencies from electrodes TP7/8, P7/8, P5/6, PO7/8, P3/4, PO3/4, O1/O2, POz and Oz. Error bars with standard error.