Neural responses to facial attractiveness: Event-related potentials differentiate between salience and valence effects

Hans Revers [a,*], Katrijn Van Deun [b], Jean Vroomen [a], Marcel Bastiaansen [a,c]

[a] Department of Cognitive Neuropsychology, Tilburg University, The Netherlands
[b] Department of Methodology & Statistics, Tilburg University, The Netherlands
[c] Academy for Leisure & Events, Breda University of Applied Sciences, Breda, The Netherlands

ARTICLE INFO

ABSTRACT

We examined the neural correlates of facial attractiveness by presenting pictures of male or female faces (neutral expression) with low/intermediate/high attractiveness to 48 male or female participants while recording their electroencephalogram (EEG). Subjective attractiveness ratings were used to determine the 10% highest, 10% middlemost, and 10% lowest rated faces for each individual participant to allow for high contrast comparisons. These were then split into preferred and dispreferred gender categories. ERP components P1, N1, P2, N2, early posterior negativity (EPN), P300 and late positive potential (LPP) (up until 3000 ms post-stimulus), and the face specific N170 were analysed. A salience effect (attractive/unattractive > intermediate) in an early LPP interval (450–850 ms) and a long-lasting valence related effect (attractive > unattractive) in a late LPP interval (1000–3000 ms) were elicited by the preferred gender faces but not by the dispreferred gender faces. Multivariate pattern analysis (MVPA)-classifications on whole-brain single-trial EEG patterns further confirmed these salience and valence effects. It is concluded that, facial attractiveness elicits neural responses that are indicative of valenced experiences, but only if these faces are considered relevant. These experiences take time to develop and last well beyond the interval that is commonly explored.

1. Introduction

Attractiveness has a major impact on how people are perceived and treated. Attractive people are judged and treated more positively than unattractive people (Langlois et al., 2000). They are, for instance, perceived as more intelligent, intellectually competent, sociable, trustworthy, to be more competent spouses, lead happier lives, and as having a more positive emotional expression than unattractive people (Feingold, 1992; Golle et al., 2014; Dion et al., 1972; Jackson et al., 1995; Ma et al., 2015; Ma et al., 2016). It is therefore of interest to explore the neural basis of the perception of attractiveness. One way to do so is by studying the event-related potentials (ERPs) to faces that differ in terms of attractiveness. However, extant ERP research on facial attractiveness has produced inconsistent and confusing results. In the present study we aim to provide clarity on this matter by measuring ERPs to maximally distinct attractiveness categories in a large sample of participants.

1.1. Attractiveness related ERP components

Through a search on Web of Science with keywords: "attractiveness and (ERP or EEG) and (face or facial)" in January 2023, we identified 18 studies that reported on ERP effects to images of faces with varying levels of attractiveness and minimal emotional expression (Table 1).

Differentiating effects in early ERP components (<300 ms) have been reported infrequently and inconsistently (see Table 1). Effects in the P2 component are reported more often than others, but in opposite directions. Effects on the early posterior negativity (EPN) component have been reported as greater amplitudes to attractive compared to unattractive faces by a few studies (Marzi & Viggiano, 2010; Werheid et al., 2007; Wiese et al., 2014).

Labelling of the observed ERP components is notoriously diverse with sometimes debatable appropriateness, especially in the late (>300 ms) interval. In an attempt to bring consistency in the labelling of ERP components, in the present work we refer to the P300 component if a component's peak latency (between 300 and 500 ms) and topography were fitting. The remaining late effects are referred to as late positive

---

Table 1

Listing of studies that reported differential ERP responses to varying levels of facial attractiveness, including the latency range and the direction of the effect.

ERP component	Study	Effect
P1 (50–120 ms)	Halit et al. (2000)	U>A
P1 (100–140 ms)	Zhang and Deng (2012)	A>U
P2 (120–220 ms)	van Hooff et al. (2011)	A/U>M
P2 (130–170 ms)	Chen et al. (2012)	U>A
P2[1] (~150 ms)	Schacht et al. (2008)	A/U>M
P2[1] (150–200 ms)	Marzi and Viggiano (2010)	A>U
P2 (150–230 ms)	Zhang and Deng (2012)	A>U
P2 (200–240 ms)	Munoz and Martin-Loeches (2015)	U>A/M
N170 (120–175 ms)	Halit et al. (2000)	U>A
N170 (140–190 ms)	Zhang and Deng (2012)	A>U
N170 (150–200 ms)	Marzi and Viggiano (2010)	A>U
N170 (150–225 ms)	Trujillo et al. (2014)	U>A
N2 (180–230 ms)	Chen et al. (2012)	A>U
N2 (180–280 ms)	Ma, Zhang, et al. (2017)	U>A
N2 (240–280 ms)	Ma et al. (2015)	U>A
EPN (200–300 ms)	Marzi and Viggiano (2010)	A>U/M
EPN (200–300 ms)	Tanaka (2021)	A>U/M
EPN (230–280 ms)	Werheid et al. (2007)	A>U
EPN (270–400 ms)	Wiese et al. (2014)	A>U
N300 (230–330 ms)	Zhang and Deng (2012)	A>U
N300 (250–350 ms)	Zhang et al. (2011)	A>U
P300 (280–420 ms)	Munoz and Martin-Loeches (2015)	A>U/M
P300[2] (300–500 ms)	Schacht et al. (2008)	A>M
P300[3] (300–500 ms)	Marzi and Viggiano (2010)	A>U
P300 (300–500 ms)	Ma, Zhang, et al. (2017)	A>U
P300[4] (330–500 ms)	Zhang and Deng (2012)	A>U
P300[5] (350–550 ms)	Zhang et al. (2011)	A>U
P300[3] (350–550 ms)	Ma et al. (2015)	A>U
P300[3] (380–480 ms)	Ma, Qian, et al., 2017	A>U
P300[5] (400–480 ms)	Roye et al. (2008)	A>U
P300 (450–650 ms)	Oliver-Rodriguez et al. (1999)	A>U correlation
LPP[6] (200–650 ms)	van Hooff et al. (2011)	A>U/M
LPP[2] (400–600 ms)	Werheid et al. (2007)	A>U
LPP (420–580 ms)	Ma, Zhang, et al. (2017)	A>U
LPP[2] (450–650 ms)	Chen et al. (2012)	U>A
LPP[2] (500–700 ms)	Munoz and Martin-Loeches (2015)	A/U > M
LPP[2] (520–1200 ms)	Roye et al. (2008)	U>A
LPP[2] (~550 ms)	Johnston and Oliver-Rodriguez (1997)	A>U correlation
LPP[2] (500–800 ms)	Schacht et al. (2008)	A/U>M
LPP (500–700 ms)	Marzi and Viggiano (2010)	A/U>M

Note. The direction of the effect is listed in abbreviated form with letters (A=attractive, U=unattractive, M=intermediate). A/U>M means that the ERP component has a larger amplitude to both the attractive and unattractive faces compared to the unattractive faces. Effects are grouped by ERP components in order of starting latency. Latencies have been rounded to the nearest 5 ms mark for clarity. Marked ERPs were identified by the authors as [1]early component/early frontal positivity, [2]late positive complex (LPC), [3]late positive potential (LPP), [4]P3b, [5]unnamed, and [6]slow wave

potential (LPP) (see Table 1).

The most frequent and consistent ERP modulations to facial attractiveness are those in the P300 with larger amplitudes for attractive compared to unattractive [A>U] and/or intermediate attractive faces [A>M] (Roye et al., 2008; Marzi & Viggiano, 2010; Zhang & Deng, 2012; Zhang et al., 2011; Oliver-Rodriguez et al., 1999; Ma, Qian, et al., 2017; Munoz & Martin-Loeches, 2015; Schacht et al., 2008; Ma, Zhang, et al., 2017; Ma et al., 2015).

Effects on LPP amplitudes (500–800 ms) have been reported frequently but in multiple directions. Some studies find larger amplitudes for attractive and unattractive compared to intermediate attractive faces [A/U>M] (Schacht et al., 2008; Marzi & Viggiano, 2010; Munoz & Martin-Loeches, 2015). However, larger amplitudes for attractive compared to unattractive faces [A>U] (van Hooff et al., 2011; Werheid et al., 2007; Ma, Zhang, et al., 2017; Johnston & Oliver-Rodriguez, 1997, or the opposite [U>A] (Chen et al., 2012; Roye

et al., 2008) have also been reported. Therefore, no conclusion can be drawn from the literature regarding LPP sensitivity to attractiveness.

The overall pattern of ERP responses to facial attractiveness does not allow for clear conclusions. Inconsistencies are likely due to differences in experimental setup, stimulus materials, and statistical methods. Fitting their specific research question, some previous studies have omitted an intermediate attractiveness category (Chen et al., 2012; Halit et al., 2000; Ma, Qian, et al., 2017; Ma et al., 2015; Ma, Zhang, et al., 2017; Roye et al., 2008; Werheid et al., 2007; Wiese et al., 2014; Zhang et al., 2011; Zhang & Deng, 2012; Trujillo et al., 2014), others have tested only participants of one gender, showing faces of the opposite gender (Ma, Qian, et al., 2017; Ma, Zhang, et al., 2017; Ma et al., 2015). Some studies presented each face multiple (Chen et al., 2012; Halit et al., 2000; Marzi & Viggiano, 2010; Munoz & Martin-Loeches, 2015; Roye et al., 2008; Schacht et al., 2008; van Hooff et al., 2011; Werheid et al., 2007; Johnston & Oliver-Rodriguez, 1997; Oliver-Rodriguez et al., 1999) or even many times (Trujillo et al., 2014; Zhang & Deng, 2012). Some have used a task related to attention (van Hooff et al., 2011), cooperation (Chen et al., 2012), fairness (Ma, Qian, et al., 2017; Ma et al., 2015), oddball detection (Zhang & Deng, 2012), or remembering (Wiese et al., 2014). While these design features may have properly served the particular purpose of specific studies and the hypotheses they were addressing, they hinder drawing more general inferences about ERP responses purely to facial attractiveness.

1.2. Current study

With the current study we aim to contribute to the field by employing a straightforward experimental paradigm, a large sample, and state-of-the-art analysis techniques in order to help us obtain more robust results. First, most of the previous studies have tested fewer than 20 participants, which was once the norm. However, the current view is that larger sample sizes are needed for quantitative electroencephalogram (EEG) studies (Höller, 2021) to have sufficient statistical power.

Second, Schacht et al. (2008) demonstrated that ERPs to rather attractive and rather unattractive faces were very similar to those of intermediate attractive faces (see Fig. 4A in their paper), pointing to the need for sharper contrasts in attractiveness. Previous studies averaged ERP responses across all trials within a pre-defined category, regardless of subjective preferences, and reported mean attractiveness ratings per category that were far from the extremes.

Third, ERP effects to attractiveness later than 800 ms seem to be largely unexplored. A visual inspection of the LPP (300–1000 ms) observed in several studies (e.g. Schacht et al., 2008; Wiese et al., 2014; Zhang et al., 2011; Oliver-Rodriguez et al., 1999) suggests that differences between attractiveness levels may extend to well beyond 1000 ms. This formed the impetus to explore latencies beyond 1000 ms in the current study.

Fourth, gender preference is a potentially decisive factor in neural response to faces. By preferred gender we mean the gender that the participant would be sexually attracted to, based on the participant's own gender and sexual orientation. Some attractiveness studies shown only male faces to heterosexual female participants (Ma, Qian, et al., 2017) or used gender as an analysis factor for only heterosexual participants (Zhang & Deng, 2012). However, sexual orientation is commonly not taken into account and the effect of gender preference on neural responses to attractiveness remains unclear.

Finally, most previous studies have used repeated measures Analysis of Variance (ANOVA) on average ERP amplitudes over predetermined time-intervals to evaluate differential responses. Analysis of time-series data like EEG using repeated measures ANOVA is not optimal due to the massive multiple comparisons problem it involves. Researchers often circumvent this issue by averaging over a latency interval and over a subset of electrodes, but this is suboptimal from a statistical point of view, as it entails a loss of variance, and promotes a strategy of double-dipping. Nonparametric, cluster-based random permutation analysis

provides an elegant method to solve the multiple comparisons problem (Maris & Oostenveld, 2007) while still allowing for incorporating prior knowledge to maximize sensitivity to the hypothesized effect.

We adopt a balanced experimental design to allow for straightforward interpretation of effects. We presented 240 images of unfamiliar male and female faces (120 each) that covered a wide range of attractiveness, from very attractive to very unattractive, having minimal emotional expression to prevent confounding factors of emotion recognition or emotion contagion. Each image was shown only once, to prevent repetition or recognition effects (Rugg et al., 1988). For example, faces are perceived as more attractive with repeated exposure, which is then also reflected in certain ERP components (Han et al., 2020). Participants gave subjective attractiveness ratings following each image on a continuous scale. We then selected the highest rated 10%, the lowest rated 10%, and middlemost rated 10% of images to form the attractive, unattractive, and intermediate attractive categories. This selection was made for each participant separately, based on their subjective ratings, to obtain an optimal attractiveness contrast. Each of the categories was then split into preferred and dispreferred gender categories based on the gender and sexual orientation of the participant. Finally, we analysed data of 48 participants, consisting of approximately equal numbers of male and female participants to avoid gender-dependent effects, using informed random permutation statistics as well as a naïve, data-driven classification approach.

We expected amplitudes in the P300 component to be greater for attractive compared to unattractive and intermediate attractive faces [A>U/M]. No clear expectations for the other ERP components, or of the later latencies (>800 ms) could be deduced from existing literature.

2. Method

2.1. Participants

Eighty healthy participants (35 male, 45 female; mean age = 20.6, SD = 2.4, range = 18 – 30 years; 65 heterosexual, 7 homosexual, 7 bisexual, 1 unknown) took part in the study after giving written informed consent. Participants were mostly first year psychology students who signed up through the university's experiment participation system and receive course credits. Participants with a bisexual orientation (n = 7) were excluded from analysis because no dispreferred gender category could be determined. Sixteen participants were excluded from analysis due to overall insufficient quality of the EEG signal. Nine additional participants were excluded because one of the conditions (see Design section below) had no artefact-free EEG-data. The remaining 48 participants (24 male, 24 female; 44 heterosexual, 4 homosexual; mean age = 20.8, SD = 2.6, range = 18 – 30 years) were included in the final analyses.

All experimental procedures were approved by the Ethics Review Board of the School of Social and Behavioural Sciences of Tilburg University (EC-2016.48).

2.2. Stimuli

In total, 252 colour images of faces were presented on a 24.5-inch BenQ Zowie XL2540 LCD screen with a resolution of 1920 × 1080 and a refresh rate of 240 Hz, against a grey background. Twelve of these were used for practice trials. The images were 250 pixels wide and 312 pixels high. On screen they were 71 by 88 mm.

These images were the result of an extensive selection and validation procedure. We first collected 3000 face images from various internet sources. From these, 400 faces were selected that were forward looking, have minimal emotion expression, minimal make-up, no tattoos or earrings. They were then cropped to show only the face. The minimal remaining background was neutral grey. According to the subjective assessment of 6 researchers/research-assistants, these selected faces covered the full attractiveness range, from very attractive to very unattractive.

The pre-selected images were then rated on attractiveness by 110 participants (89 female, 21 male; mean age = 20.5, SD = 3.7) who did not overlap with the EEG study, through an online Qualtrics (https://www.qualtrics.com) questionnaire. With these ratings, the images were categorized into female/male × attractive/ intermediate/unattractive attractive faces, each containing 40 images for the experiment and 2 for the practice session. Covering a hiatus in attractive male faces, we complemented this category with 10 images from an existing dataset (Pronk & Denissen, 2020) that adhered to the same criteria and were edited in the same manner as the already selected images.

We then collected ratings of emotional expressions of the stimuli in a lab setting in a pilot study, from 44 participants (39 female, 5 male; mean age = 20.0, SD = 2.2) who did not participate in either the online behavioural or the EEG study. The attractive, intermediate, and unattractive faces were rated on a 7-point continuous scale (0 = neutral) as having minimal emotion expression, with a mean (SD) of 0.09 (0.2), 0.06 (0.2), and 0.7 (0.3), respectively.

Finally, we assessed whether luminance and contrast of the images were similar between pre-defined attractiveness categories through MATLAB (v. R2019a, MathWorks, Inc.). Each image was first converted to the perceptually realistic CIELAB (or L*a*b*) colour space (ISO/CIE, 2019). Independent-sample t-tests showed that luminance was similar between the attractive (M = 70.7, SD = 5.9) and intermediate faces (M = 71.0, SD = 6.2), $t(158)$ = −0.37, p = .71. The unattractive faces (M = 73.7, SD = 5.1) had significantly higher luminance compared to the attractive faces, $t(158)$ = −3.45, p < .001, and the intermediate faces, $t(158)$ = 2.96, p = .004. However, in absolute sense, these differences in luminance were very small; 3 points or less on a 100 point scale. Average contrast was similar between attractive (M = 16.4, SD = 4.1), intermediate (M = 16.3, SD = 4.2), and unattractive faces (M = 15.3, SD = 3.7), all $t(158)$-values ≤ 1.74, all p-values ≥ 0.09.

2.3. Design

Following a 1000 ms fixation cross, a face images was presented for 1000 ms, then a 2000 ms blank, grey screen, and finally a horizontal slider that could be moved on a continuous scale to rate attractiveness (Fig. 1). The slider always started in the middle (value = 0). The far left was labelled with "VERY UNATTRACTIVE" (value = −3) and the far right was labelled "VERY ATTRACTIVE" (value = 3). Values were not shown. Below the slider was a button labelled Next to continue to the next trial. Participants were instructed that the button had to be pressed within 5000 ms in order for the rating to be recorded and to use their "gut-feeling".

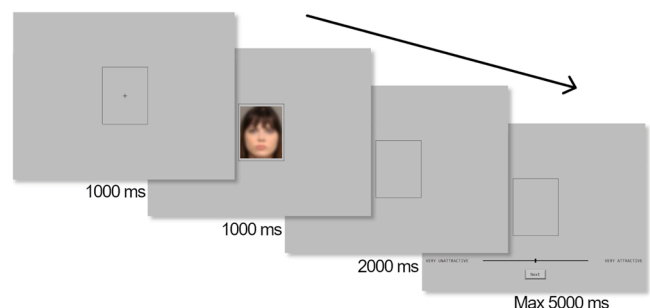Each participant viewed all 240 images in a within-subjects design.

Fig. 1. Graphical representation of a trial sequence. Note: The sequence of a single trial: A 1000 ms fixation cross, the stimulus (image of a face) for 1000 ms, a blank response delay of 2000 ms, and finally a response slider that would proceed to the next trial when the "Next" button was pressed or when 5000 ms had passed. For the slider, the label on the far left read "VERY UNATTRACTIVE", the label on the far right read "VERY ATTRACTIVE". The face image was blurred for this publication only. The faces that were shown in the experiment were not blurred.

The order was semi-randomized, such that each block of 24 trials consisted of 4 random images from each of the 6 predefined categories (female/male × attractive/ intermediate/unattractive attractive) in randomized order. Each image was shown only once.

Subjective attractiveness ratings were used as independent variables. For exploring the contrasts between neural responses to attractive, intermediate, and unattractive faces, we selected the 24 images (10%) with the highest, 24 images with the lowest and 24 images with the middlemost ratings, based on the subjective ratings of each participant individually. These selections were used in order to compare the most pronounced differences and will from now on be referred to as HI, LO, and MID, respectively. These categories were further split into preferred gender and dispreferred gender subcategories according to the combination of participant gender, gender of the face stimulus, and sexual orientation of the participant. Dependent variables are the EEG amplitude for the ERP analysis of the 3.0 s interval following stimulus onset.

2.4. Procedure

After reading information regarding the experiment and giving informed consent in accordance with the Declaration of Helsinki, the participants were asked to wash their hands with lukewarm water, without soap, as is recommended for the BioSemi skin conductance measurements. The participant was then prepared for the physiological measurements, which took approximately 30 min. Following the preparation, the participant was seated in a dimly lit, sound attenuating cabin, approximately 60 cm from a computer screen. The experiment was carried out with E-Prime 3.0 software (Psychology Software Tools, Pittsburgh, PA) and started with 12 practice trials to get familiar with the procedure and the rating slider, followed by the actual experiment consisting of 240 trials. After each block of 24 images, a self-paced short break was offered. Viewing all images took between 40 and 50 min. At the end, we asked the participant to indicate their sexual preference (heterosexual, homosexual, bisexual, other, or prefer not to answer).

2.5. Physiological measurements

EEG was recorded with 64 Ag-AgCl ActiveTwo electrodes from BioSemi in an extended 10/20 layout (Chatrian et al., 1985) at a sampling rate of 512 Hz. Impedance was kept below 5 kΩ during recording. SignaGel was used to facilitate conduction. Two electrodes were applied on the mastoids, behind the ears, for offline re-referencing. For detection of eye blinks and eye movements, we applied one electrode above, and one below the right eye, and an electrode next to each of the outer canthi of the eyes.

Additionally, we measured electrical activity from three facial muscles (fEMG): the Zygomaticus Major, the Orbicularis Oculi, and the Corrugator Supercilii. Electrodermal activity (skin conductance) was measures through two BioSemi GSR electrodes attached to the distal phalange of the index and middle finger of the left hand. Skin conductance and fEMG data was not analysed for this paper.

2.6. Preprocessing

The raw EEG data was preprocessed in BrainVision Analyzer 2.1.2 (Brain Products GmbH). First, the data was re-referenced to the mean of the mastoid channels. Slow drift and high frequency noise were filtered out using Zero phase shift Butterworth filters (0.01 Hz high-pass and 100 Hz low-pass, −12 dB/octave roll-off). The signal was then segmented into epochs from 700 ms before until 3200 ms after stimulus onset. Baseline correction was applied using the 200 ms interval leading up to stimulus onset as the baseline interval. Channels with overall poor signal quality were reconstructed by fourth order spherical splines interpolation. The number of reconstructed channels was limited to 10%. On average 1.8 channels were reconstructed per participant. Next, artefacts caused by eye blinks and eye movements, and others, were

corrected using Independent Component Analysis (ICA). Prior to ICA, major artefacts were manually marked as bad segment on individual channels. The remaining data was used for the decomposition of the ICA procedure. Ocular components were manually identified and excluded in the inverse ICA process. Through semi-automatic inspection, trials with blinks or eye movements at stimulus onset were rejected, as were any remaining artefacts. Finally, the EEG was again baseline corrected using the 200 ms before stimulus onset as baseline. Participants were excluded from further analysis if their preprocessed EEG data did not meet predetermined criteria of having no more than 10% (>6) bad channels and no more than 50% of the segments removed due to artefacts. The remaining participants had, on average, 200.6 of the 240 trials left after artefact rejection.

The N170 and EPN components are commonly analysed using a Common Average Reference (the average of all electrodes) as these components are topographically close to the mastoids. Thus, for the purposes of inspecting the N170 and EPN components, the data was separately processed using a Common Average Reference.

Further processing steps and analyses were performed in MATLAB (v. R2019a, MathWorks, Inc.) using the FieldTrip Toolbox (Oostenveld et al., 2011). The data was low-pass filtered at 30 Hz prior to ERP analysis.

2.7. Statistical analysis

To compare neural responses to seeing attractive vs intermediate vs unattractive faces, we first constructed clearly distinct categories by selecting trials corresponding to the highest (HI), lowest (LO), and middle (MID) 10% of the subjective attractiveness ratings for each participant individually. All statistical analyses were done in MATLAB (v. R2019a, MathWorks, Inc.) and the FieldTrip Toolbox (Oostenveld et al., 2011).

2.7.1. Cluster analysis
Nonparametric cluster-based permutation tests (Maris & Oostenveld, 2007) were used to determine statistical significance of the differences between HI, MID, and LO. Multiple comparison correction was applied on the cluster level. For the sake of readability, only significant *p*-values are provided.

First, we specifically tested for differences between attractiveness categories in ERP components P1 [80–100], N1 [80 − 120], P2 [130 − 160], N2 [270 − 320], EPN [220 − 255], P300 [300–400], and LPP [300 − 1000 ms] in their respective, commonly reported latency intervals (Hajcak et al., 2011) using all 64 channels. Additionally, we tested for differences in the face-specific N170 [130 − 165 ms] using only the occipito-temporal electrodes (T7/8, TP7/8, P7/8/9/10, PO7/8, O1/2) as advised by Rossion and Jacques (2008).

In a subsequent exploratory approach, we performed a cluster analysis on the 3000 ms interval following stimulus onset without averaging over time to obtain clusters irrespective of specific ERP components.

2.7.2. Classification
Additionally, as a data-driven approach to complement the theory-driven inferential statistics, we performed multivariate pattern analyses (MVPA). Decoding techniques like MVPA have become standard practice in the analysis of fMRI data. The technique has also been increasingly applied to EEG data in the past decade (Grootswagers et al., 2017). We have performed MVPA classification with the MVPA-light toolbox for MATLAB (Treder, 2020) as implemented in the FieldTrip toolbox (Oostenveld et al., 2011) using a Linear Discriminant Analysis (LDA) classifier.

MVPA was done on the ERP amplitudes of the pre-processed whole-scalp EEG of the 3000 ms interval after stimulus onset, to determine whether successful classification could be performed on the trial level. The EEG patterns of the HI, MID, and LO trials were classified with a

single-trial classification procedure. In order to obtain a generalizable classification model, rather than separate participant-specific ones, we combined all trials of all participants which were then treated as a single dataset. This approach increases the generalizability of the model (Frid & Manevitz, 2020), as this single model was able to classify data from all participants in the current study. This arguably is of greater value to advance knowledge of neural responses to attractiveness than many separate (and different) models that are optimized to each specific individual participant. Additionally, due to the large amount of trials per condition (12 trials x 48 participants = 576 maximum, minus artefact rejected trials) in this combined dataset, the classification results would be highly reliable.

To prevent overfitting, we used a 5-fold validation procedure in which the data was randomly divided in 5 parts. The classifier was then trained on 4 parts, and the resulting model was tested on the fifth. This procedure is repeated so that each part functions as test data once. This cross-validation procedure is repeated 5 times with a new random division of the data into 5 folds. Stratified sampling was used to create the folds so that class proportions were approximately preserved.

Our objectives were to explore whether it would be possible to decode the self-reported attractiveness category of the viewed face based on single-trail data. For this, we first determined the accuracy per time-point across electrodes to explore at which latency intervals the classification model leads to the most accurate decoding. Additionally, we determined the accuracy per time-point, per electrode to explore the development of decoding accuracy over space and time. These analyses were performed both on the three class model, classifying between the HI, MID, and LO classes to obtain the general decoding accuracy, and on each of the pairs of classes to allow for pairwise comparisons and so determine the presence of emotional salience and valence effects.

We also assessed pairwise classification accuracy across time for the latency intervals with the highest classification accuracy levels to allow for comparison of the topographical maps, both between intervals and between the pairwise classifications.

3. Results

3.1. Attractiveness categories

We selected the trials with the highest (HI), lowest (LO), and middle (MID) ratings per participant, 24 images (10%) in each category. These were then split into preferred gender and dispreferred gender subcategories. Of the preferred gender, the HI images ($M = 1.68$, $SD = 0.55$) were rated higher ($t(47) = 25.22$, $p < .001$) than the MID images ($M = -0.22$, $SD = 0.48$), and higher ($t(47) = 36.27$, $p < .001$) than the LO images ($M = -2.44$, $SD = 0.42$). MID images were also rated higher ($t(47) = 26.99$, $p < .001$) than LO images (Fig. 2, blue bars).

Of the dispreferred gender, the HI images ($M = 1.57$, $SD = 0.56$) were rated higher ($t(47) = 25.48$, $p < .001$) than the MID images ($M = -0.26$, $SD = 0.48$), and higher ($t(47) = 36.10$, $p < .001$) than the LO images ($M = -2.49$, $SD = 0.43$). MID images were also rated higher ($t(47) = 27.49$, $p < .001$) than LO images (Fig. 2, red bars).

The attractiveness categories are clearly distinct for both the preferred and dispreferred gender in terms of subjective attractiveness ratings, providing a thorough basis for exploring the differences in neural responses to differences in perceived attractiveness.

The ratings for the preferred gender faces were slightly higher than from those of the dispreferred gender faces for the HI ($t(47) = 4.87$, $p < .001$), MID ($t(47) = 3.71$, $p < .001$), and LO categories ($t(47) = 3.36$, $p = .002$), however, these differences were very small (0.11, 0.04, and 0.05 rating points, respectively).

Division between preferred and dispreferred was approximately equal, with slightly more preferred gender images in each of the HI, MID, and LO categories (Fig. 3).
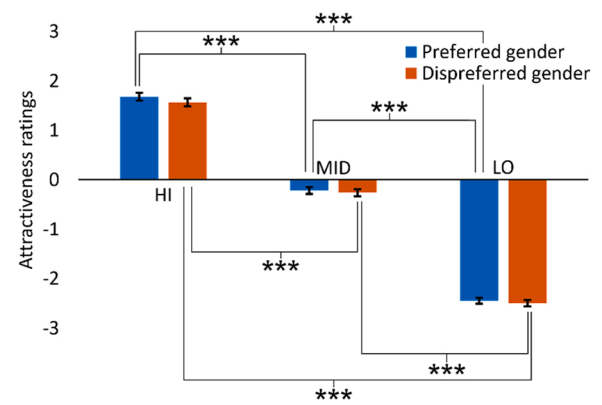

Fig. 2. Mean attractiveness ratings per category. Note. The mean attractiveness rating for the categories of the 10% highest (HI), middle (MID), and lowest rated (LO) images on the participant level of the preferred (blue bars) and dispreferred (red bars) gender. Error bars represent the standard error of the mean. Significant differences are indicated between attractiveness categories only. ***$p < .001$.
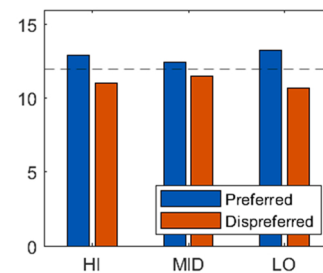

Fig. 3. Average number of preferred and dispreferred faces in each category. Note. Each of the HI, MID, and LO categories consists of 24 faces. The dashed horizontal line at 12 is added as the reference for which the number of preferred and dispreferred faces would be equal.

3.2. Event related potentials

Differences in ERPs between each of the HI, MID, LO categories were explored using cluster analysis on specific latency intervals. In addition, we performed exploratory cluster analysis on the 3000 ms EEG data post stimulus onset, without averaging over time, to detect clusters of significant differences between attractiveness categories, unbiased by existing literature.

3.2.1. Analysis of ERP components

3.2.1.1. Preferred gender. Cluster analysis of differences in specific predetermined ERP intervals between the HI, LO, and MID categories of the preferred gender faces revealed that the P300 component was significantly larger for the HI compared to the MID ($p = .027$) category in the frontal-central, parietal, midline area (Fig. 4, middle column). There were no clusters of significant differences between the LO compared to the HI or MID category.

The LPP component was significantly larger for the HI compared to the MID ($p < .001$) and to the LO categories ($p = .032$), and larger for the LO compared to the MID ($p = .001$) (Fig. 5A).

For the preferred gender categories, the data contained no significant differences for any of the other analysed ERP components: P1, N1, P2, N2, EPN, and N170 (Fig. 4, middle).

3.2.1.2. Dispreferred gender. For the dispreferred gender categories, the P1 component was larger for the LO compared to the MID category ($p = .034$) in left-temporal-parietal regions (Fig. 4, right column). The
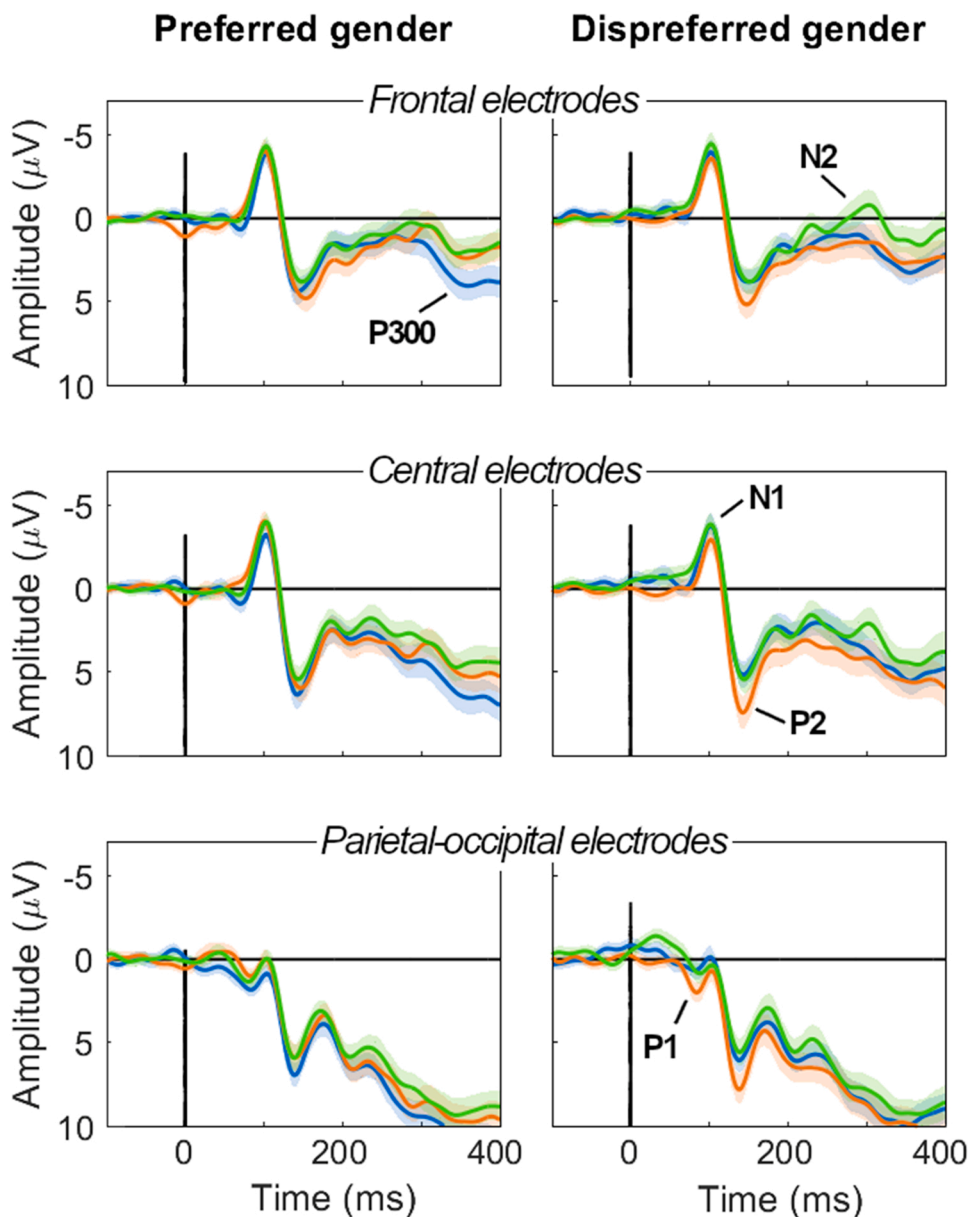
Fig. 4. Early ERPs to the 10% highest, middlemost, and lowest rated faces. Note. The averaged EEG signals of the 10% highest (HI; blue), middlemost (MID; green), and lowest (LO; orange) rated faces for images of the preferred gender (left) and the dispreferred gender (right). The EEG signals is averaged over frontal (top row; Fp1/2/z, AF3/4/7/8/z, F1/2/3/4/5/6/7/8/z), central (second row; FC1/2/3/4/z, C1/2/3/4/z, CP1/2/3/4/z), and parietal-occipital sites (third row; P1/2/3/4/z, PO3/4/z, O1/2/z). Shaded areas around the EEG represent the standard error or the mean.

HI category did not differ significantly in P1 amplitude from either the LO or the MID category.

The N1 component had a significantly larger negative deflection for the MID compared to the LO dispreferred-gender category ($p = .046$) in right-central-frontal regions (Fig. 4, right). The HI category did not differ significantly from the LO and MID categories.

Amplitudes of the P2 component were significantly larger to the LO compared to the MID ($p = .005$) and to the HI dispreferred-gender category ($p = .017$) in central regions (Fig. 4, right). Amplitudes did not differ significantly between the HI and MID categories.

The N2 component had a significantly larger negative deflection to the MID compared to the LO category ($p = .016$) in frontal regions (Fig. 4, right). N2 amplitudes did not differ significantly between the HI and MID, and between the HI and LO categories.

For the dispreferred gender faces, the P300 component did not differ significantly between any of the HI, MID, and low categories. The LPP component was significantly larger for the HI compared to the MID ($p < .040$), and larger for the LO compared to the MID ($p = .001$) (Fig. 5B). The LPP did not differ significantly between the HI and LO categories.

No significant differences between dispreferred-gender categories were observed in the N170 and EPN components.

3.2.2. Exploratory analysis of the 0 – 3000 ms interval

3.2.2.1. Preferred gender. Exploratory cluster analysis of ERP difference between the HI, LO, and MID categories of the preferred gender faces in the 0–3000 ms interval revealed clusters of significant differences in each of the pairwise comparisons. These clusters all lie either within the interval typically associated with the LPP component or beyond the time scope that is usually explored. For lack of existing terminology we will refer to clusters in the typical LPP interval as the early LPP, and clusters in the later interval as the late LPP, as these two are clearly distinguishable.

Amplitudes of the responses in both the HI ($p < .001$) and the LO ($p = .006$) categories were larger compared to those in the MID category in central-parietal regions from 340 until 1551 ms and from 430 until 867 ms, respectively. We identify these as early LPP effects (Fig. 5A) and note that the pattern of responses corresponds to a salience effect: larger amplitudes to the positive and negative compared to the neutral stimuli.
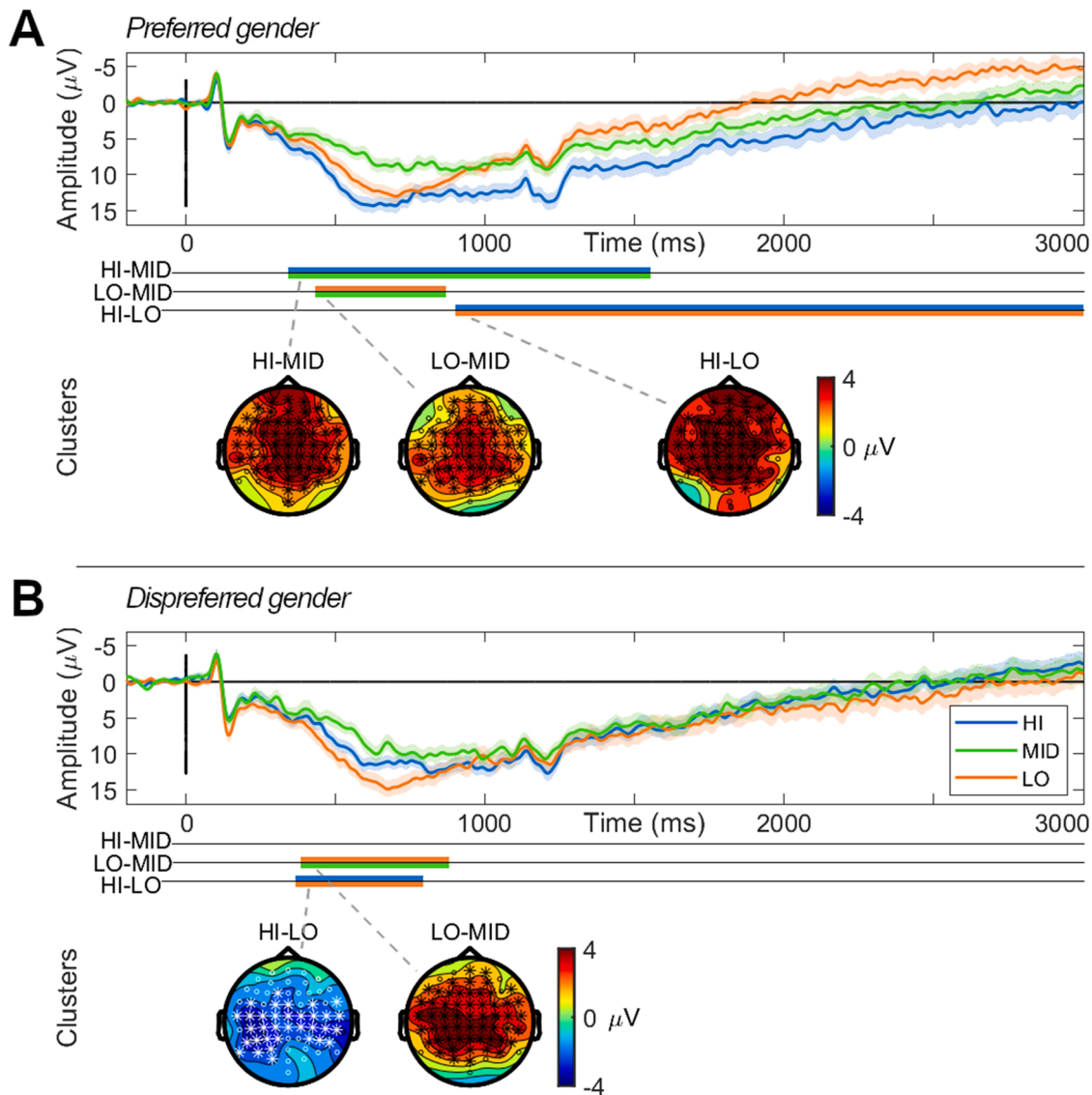
Fig. 5. EEG amplitudes and clusters of differences between the HI, MID, and LO categories. Note. A. The averaged EEG signals over central (FC1/2/3/4/z, C1/2/3/4/z, CP1/2/3/4/z) electrodes for the 10% highest rated (HI; blue), middlemost rated (MID; green), and lowest rated (LO; orange) faces of the preferred gender. Shaded areas around the EEG represent the standard error or the mean. Peaks in the 1.1–1.3 s interval are offset potentials as the stimulus was removed after 1 s. The screen remained blank (grey) for the 2 following seconds. Coloured horizontal lines beneath the EEG show the time-intervals of the clusters of significant differences between each of the HI, MID, and LO categories, with the topographies of each cluster. Colours represent the average difference in amplitude between the two indicated categories. The circles mark electrode placement, asterisks mark electrodes that are part of the cluster. Panel B shows the same for the dispreferred gender.

For mean amplitudes, see Table 2.

In later intervals, amplitudes to the HI category were larger ($p = .002$) than those to the LO category in central and pre-frontal regions from 897 until 3000 ms. We identify this as a late LPP effects (Fig. 5A) and note that the pattern of responses corresponds to a valence effect: larger amplitudes to the positive compared to negative stimuli, with responses to neutral stimuli in between. The MID category (neutral stimuli) did not differ from either the HI (positive) and the LO (negative) categories. (Fig. 5B).

3.2.2.2. Dispreferred gender. Amplitudes were larger for the LO compared to the MID dispreferred gender faces ($p = .016$) in central-parietal regions from 381 until 879 ms. Amplitudes to the HI category were smaller than those in the LO category ($p = .043$) in central-parietal regions from 365 to 791 ms (Fig. 5B). There was no cluster of significant differences between the HI and MID categories in the early LPP interval, or between any two categories in the late LPP interval.

Response patterns to dispreferred gender faces show no clear salience effect in the early LPP interval and do not show a valence effect in the late LPP interval (Fig. 5B).

3.2.3. Patterns of ERP amplitudes of all ratings

To assess the reliability of these results, we additionally explored the ERP responses to each of the 10% rating bins. Again, we ordered all subjective attractiveness ratings for each individual participant and split these in 10% bins, each containing 24 trials per participant. We then split these bins into preferred and dispreferred gender bins, and aggregated the EEG waveform over trials and participants for each bin (Fig. 6). Mean amplitudes were calculated per 10% bin for the early and late LPP component, aggregating over the time-interval and over the electrodes that constitute the respective cluster (Fig. 6, bar chart insets).

For the preferred gender, the mean amplitudes per bin show patterns that are supportive of the results of the cluster analysis. The U-shaped pattern of the early LPP bins corresponds with a salience effect, and the linear trend in the late LPP corresponds with a valence effect (see the bar

Table 2
Mean amplitudes of the ERP components.

ERP component	Mean (SD) amplitude (µV)		
	HI	MID	LO
Preferred gender			
P300	5.28 (6.42)	2.97 (5.33)	3.56 (5.69)
early LPP	10.71 (4.88)	6.33 (4.97)	9.17 (5.17)
late LPP	4.38 (8.36)	2.03 (6.56)	-0.36 (6.24)
Dispreferred gender			
P1	-1.79 (3.53)	-2.13 (3.12)	-0.89 (3.68)
N1	-1.48 (3.56)	-1.92 (3.33)	-0.54 (3.61)
P2	3.97 (4.13)	3.99 (4.46)	5.84 (5.36)
N2	1.82 (4.90)	0.30 (5.57)	2.36 (6.25)
early LPP	8.82 (4.58)	7.08 (5.51)	10.74 (6.22)

Note. The mean amplitudes and standard deviations for each ERP component. Only those components that showed significant differences between (HI, MID, LO) conditions are listed, for each of the preferred and dispreferred gender categories. Means were calculated as the average amplitudes over the time-interval and over the electrodes that correspond to the respective significant cluster.

chart insets in Fig. 6, upper half). The ERP waveforms further show that these salience and valence effects are consistent over the corresponding time intervals.

For the dispreferred gender, the pattern of the amplitudes of the early LPP bins roughly resembles the U shape of a salience effect, though less

convincing then that of the preferred gender. The bins of the late LPP interval do not show a consistent pattern (bar chart insets in Fig. 6, lower half). ERP waveforms of the late LPP interval confirm there are no apparent systematic differences between bins throughout the interval.

3.3. Classification

For the MVPA classification of the preferred (and the dispreferred) categories, the trials of all participants were combined per category into 496 (451) HI trials, 518 (463) MID trials, and 533 (438) LO trials. Note that the sum of the preferred and dispreferred numbers per category is less than the maximum amount of 1152 trials (24 trial x 48 participants) due to artefact and blink rejection.

3.3.1. Preferred gender

Overall classification accuracy using all three preferred gender categories (HI, MID, and LO) per time-point, per electrode, reached accuracy levels around 38% (chance level 33%) in the time period roughly equivalent to the cluster interval of the early LPP effect (450–850 ms). Accuracy in the late LPP interval (1000–3000 ms) fluctuates around chance level with some small clusters of around 37% accuracy (Fig. 7A).

To allow for pairwise comparisons, we also classified each of the pairs of categories. Both the HI vs MID classification and the LO vs MID classification reached relatively high accuracy levels (~58% and ~57%, respectively) in the early LPP interval (Fig. 7B) while the HI vs LO
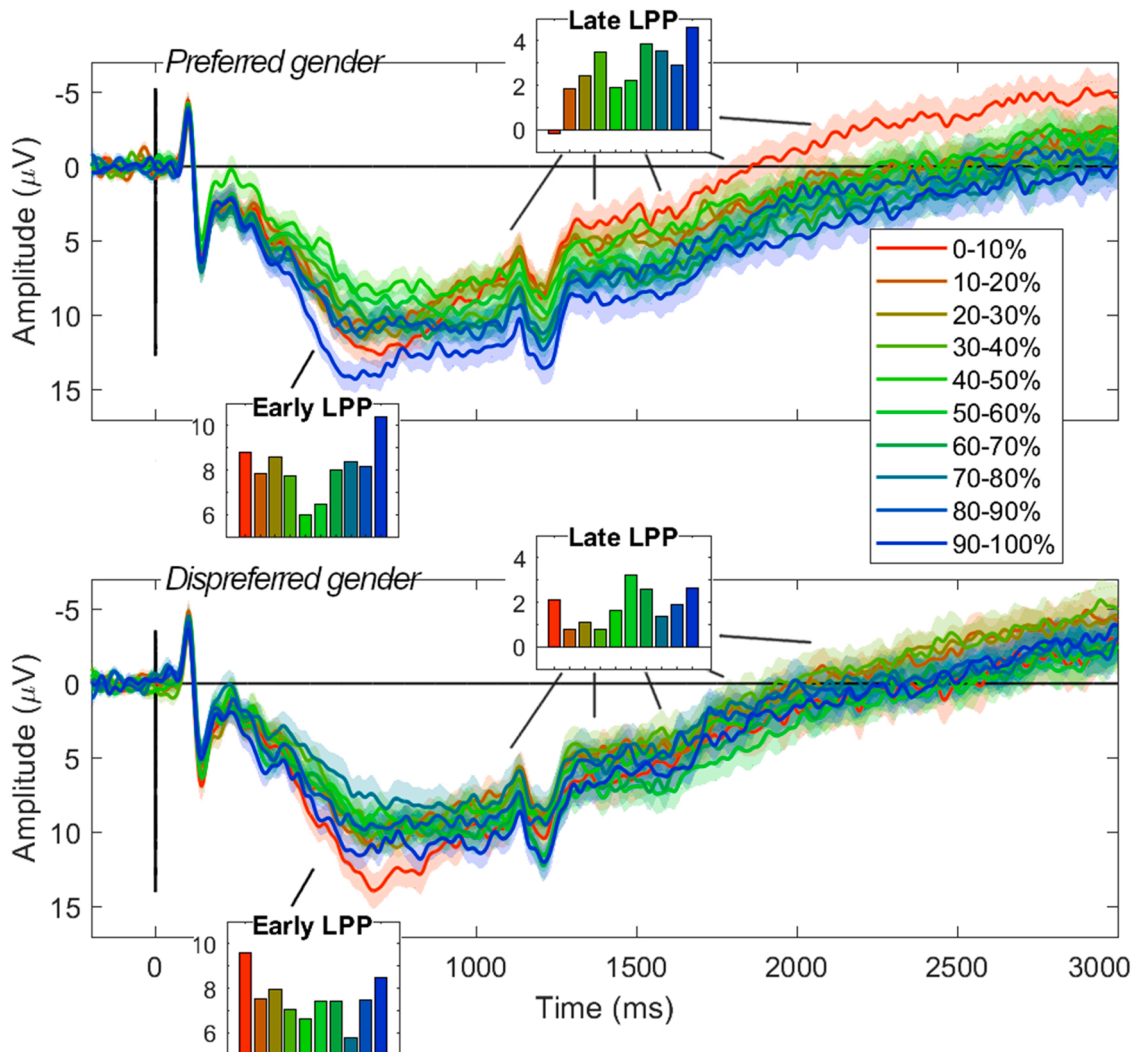

Fig. 6. ERP amplitudes of all trials split into 10% bins of the ordered ratings per participant, for the preferred and dispreferred gender.
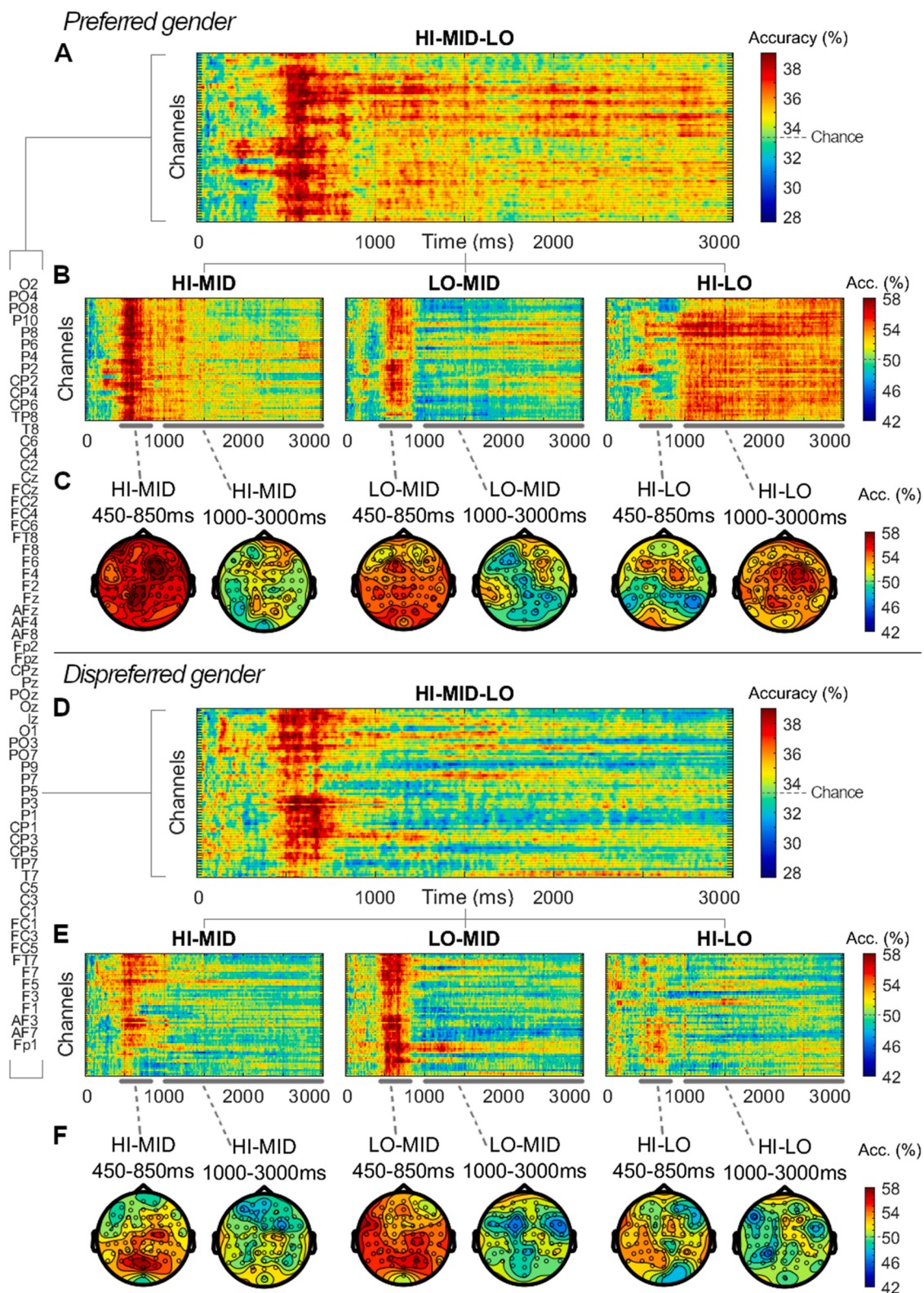
Fig. 7. Overview of the classification accuracy. Note. A. MVPA classification accuracy per time segment, per electrode of the three way (HI, MID, and LO) classification of the preferred gender data. Chance level is 33%. B. Classification accuracy per time segment, per electrode of the pairwise (HI vs MID, LO vs MID, and HI vs LO) classifications of the preferred gender data. Chance level is 50%. C. The topographical maps of classification accuracy for the early LPP interval (450–850 ms) and the late LPP interval (1000–3000 ms) of the preferred gender data as detected in the ERP analysis as reported above. D, E, and F. Like A, B, and C above, for the dispreferred gender data.

classification showed accuracies around chance level. Topographical maps of the classification accuracy averaged over the early LPP interval confirm that the classifier is modestly able to distinguish both the HI and LO from the MID category (Fig. 7C), but performs considerably less well in distinguishing between the HI and LO categories. This shows that a differentiating early LPP effect is detectable at the single trial level across participants.

A pattern of classification accuracy consistent with the late LPP effect was found for the 1000–3000 ms interval. The HI was distinguished from the LO category with moderate (54–55%), but long lasting, global classification accuracy, while both HI and LO were classified from the MID class with accuracies near chance level (Fig. 7B). This is further confirmed by topographical maps of the average accuracies over this interval (Fig. 7C), showing that a differentiating late LPP effect is also detectable at the single trial level across participants.

3.3.2. Dispreferred gender

Classification of all three dispreferred gender categories per timepoint, per electrode resulted in an overall lower accuracy then classification of the preferred gender. Accuracies of around 38% are achieved in the early LPP time period, however these accuracy peaks are more scattered and scarce than in the preferred gender classification. Accuracy fluctuates around chance level in the late LPP interval. (Fig. 7D).

Pairwise comparisons revealed that accuracy levels in the early LPP interval were relatively high in the LO vs MID classification (~57%) (Fig. 7E) but lower for the HI vs MID (~54%) and the HI vs LO classifications (~chance level). Topographical maps of the classification accuracy averaged over the early LPP interval further confirm these results (Fig. 7C). A differentiating early LPP effect is modestly detectable between the LO and the MID categories at the single trial level across participants, but not between HI and MID or between HI and LO. Importantly, the above reported cluster of significant differences between the HI and LO categories is not confirmed by single trail classification.

The pairwise classifications show only near chance level accuracies for each of the pairwise combinations in the late LPP interval (Figs. 7E and 7F). No differentiating late LPP effect is detectable at the single trial level across participants. The valence effect appears absent for the dispreferred gender trails.

4. Discussion

The aim of the present study was to provide clarity on which ERP responses are sensitive to facial attractiveness, and how these responses differ between attractiveness levels. We showed 240 images of faces to 63 participants while recording their EEG in a simple straightforward experimental setup. To obtain optimal contrasts, we made a selection per participant of the faces that were rated as the 10% most attractive, 10% least attractive, and 10% intermediate attractive based on the subjective ratings of that participant. These categories were further split into preferred and dispreferred gender categories based on the gender and sexual preference of each participant. We analyzed and compared the ERP responses to the attractive, intermediate attractive and unattractive faces of the preferred and of the dispreferred gender. Each of the ERP components of interest (P1, N1, P2, N2, N170, EPN, P300, and LPP) was clearly present in the average EEG waveforms. Differential amplitudes were observed for the P300, early and late LPP components for the preferred gender faces, and for the P1, N1, P2, N2, and early LPP components for the dispreferred gender faces. We discuss the LPP effects first since they show the most robust and interesting effects.

4.1. Early LPP interval: salience effects

ERP amplitudes in the early LPP interval (450–850 ms) to faces of the preferred gender demonstrated robust effects as a function of attractiveness. Amplitudes to both attractive and unattractive faces were greater compared to intermediate attractive faces [A/U>M] over almost the entire scalp, most pronounced over central-parietal areas. Classification gave further support to these statistical results. This pattern matches that of common findings of (early) LPP responses from valence studies (Hajcak et al., 2011) of larger ERPs to both positive and negative compared to neutral valenced stimuli, labelled as an emotional salience effect.

Similar effects have been reported in several attractiveness studies (e.g. Schacht et al., 2008; Marzi & Viggiano, 2010; Munoz & Martin-Loeches, 2015). Importantly, these studies all adopted a straightforward experimental paradigm involving attractiveness judgment, like the current study. Studies that used the faces as a distractor (van Hooff et al., 2011), or in a priming (Ma, Zhang, et al., 2017; Werheid et al., 2007) or trust paradigm (Chen et al., 2012) reported different effects (see Table 1). Possibly these paradigms distracted from the attractiveness feature of the stimuli, preventing a valenced experience. Other studies did not include an intermediate attractiveness condition (Werheid et al., 2007; Chen et al., 2012; Ma, Zhang, et al., 2017; Roye et al., 2008) precluding comparisons with a neutral valenced condition.

Early LPP responses to dispreferred gender faces, however, showed a different pattern. While responses to unattractive faces were larger than those to intermediate attractive faces, those to attractive faces were not. Additionally, responses to unattractive faces were larger than those to attractive faces, though the classification result from did not support this. Taken together, early LPP responses to the dispreferred gender faces are largely, but not entirely a salience effect.

4.2. Late LPP interval: valence effects

The usually unexplored late LPP amplitudes were greater to attractive compared to unattractive faces [A>U] of the preferred gender from approximately 1000 ms until the end of our analysis interval at 3000 ms. This long lasting effect with widespread topography, peaked at central and pre-frontal areas. Within this interval, ERP amplitudes to the intermediate attractive faces were largely in between those to the unattractive and attractive faces, with no significant difference to either from 1550 ms onward. Classification also supported these statistical results.

This pattern of responses corresponds to an orderly valence effect of larger amplitudes to pleasant stimuli (attractive faces) than to unpleasant stimuli (unattractive faces), with responses to neutral stimuli (intermediate attractive faces) in between.

Late LPP differences to the dispreferred gender faces, on the contrary, do not differentiate between attractive, intermediate attractive, and unattractive faces in the late LPP interval. Hence, facial attractiveness of the dispreferred gender does not seem to elicit valence specific responses. Classification confirmed the absence of a valence effect in the late LPP interval for the dispreferred gender faces.

The current findings extend previous reported effects to facial attractiveness in showing that ERP differences to facial attraction evaluation take time to develop. Robust effects continue well beyond the 800 ms interval. Future studies of attraction should take this observation into consideration, and structure their experimental paradigms in a way that allows for analysing longer epochs than is now common in the field.

Additionally, such late ERP effects suggest controlled, rather than automated information processing. Future studies should explore this, for instance through a second task manipulation.

The main added value of our classification lies in the specific approach of combining all single-trial data of all participants and treating those as a single dataset. This means that the classifier was trained on trials belonging to several participants, and the resulting model was then tested on different trials of several (potentially different) participants. Further, classification was performed on single trials. That we still found moderate classification accuracies that were in line with the results of the cluster analysis, despite the added variance of individual differences, speaks for the generalizability of the results.

4.3. Other ERP components

Of the remaining ERP components that we analysed (P1, N1, P2, N2, N170, EPN, and P300), only the P300 appeared sensitive, and was larger for attractive compared to intermediate attractive faces [A>M].

Interestingly, the opposite was true for ERP effects for the dispreferred gender faces: the P1, N1, P2, and N2 component showed differences between attractiveness levels, while the P300 component did not. Amplitudes of the P1 component were larger for unattractive compared to intermediate attractive faces [U>M], the N1 component had a larger negative amplitude for the intermediate compared to the unattractive faces [M>U], the P2 component was larger for the unattractive compared to the intermediate attractive and the attractive faces [U>M/A], and the N2 component had a larger negative amplitude to the unattractive compared to the intermediate attractive faces [U>M].

In comparing the current findings with those from previous studies, we can only note that the P300 effect to faces of the preferred gender are in line some attractiveness studies (Schacht et al., 2008; Munoz & Martin-Loeches, 2015) although most studies found larger P300 amplitudes to attractive compared to unattractive faces (Roye et al., 2008; Marzi & Viggiano, 2010; Zhang & Deng, 2012; Zhang et al., 2011; Oliver-Rodriguez et al., 1999; Ma, Qian, et al., 2017;; Ma, Zhang, et al., 2017; Ma et al., 2015; Munoz & Martin-Loeches, 2015). The other ERP components have been reported by some studies, but in conflicting directions.

In general, faces of the dispreferred gender evoke early (<300 ms) differential responses while faces of the preferred gender appear to evoke late (>300 ms) differential responses in patterns resembling affective salience and valence effects. A plausible interpretation is that the preferred gender faces are relevant to the participant in a manner that the dispreferred gender faces are not. This self-relevance results in a more conceptual processing of the affective value of the faces, which generally occurs at longer latencies. The absence of self-relevance of the dispreferred gender faces results in perceptual evaluation, which generally occurs at early latencies.

4.4. How does attractiveness relate to emotion?

It has been suggested before that attractive faces may induce affective responses (Olson & Marshuetz, 2005). Indeed, attractiveness ratings have been shown to correlate highly with subjective valence ratings (Oosterhof & Todorov, 2008; Todorov & Engell, 2008; Yuan et al., 2021) and responses to attractiveness are in many ways similar to affective responses like the activation of certain facial muscles (Principe & Langlois, 2011; Schein & Langlois, 2015; Gerger et al., 2011).

The current finding of a salience effect in the early LPP interval is in line with consensus from emotion research that LPP components are commonly larger for both pleasant and unpleasant compared to neutral images (Hajcak et al., 2010; Hajcak et al., 2011).

We recommend that future studies explicitly explore the exact relationship between facial attractiveness and emotions. Showing images of faces with varying levels of attractiveness would be a relatively simple, straightforward, and unambiguous emotion elicitation paradigm that could be of great value to the field of emotion research.

5. Conclusion

Facial attractiveness elicits neural responses that are indicative of valenced experiences. However, this is only the case for faces of the preferred gender, in other words, for faces that are self-relevant. Responses to facial attractiveness are clearly separated in time, suggesting a sequential process. Processing of affective salience is followed by the processing of valence. These experiences take time to develop and last well beyond the interval that is commonly explored.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.biopsycho.2023.108549.

References

Chatrian, G. E., Lettich, E., & Nelson, P. L. (1985). Ten Percent Electrode System for Topographic Studies of Spontaneous and Evoked EEG Activities. *American Journal of EEG Technology, 25*(2), 83–92. https://doi.org/10.1080/00029238.1985.11080163

Chen, J., Zhong, J., Zhang, Y., Li, P., Zhang, A., Tan, Q., & Li, H. (2012). Electrophysiological correlates of processing facial attractiveness and its influence on cooperative behavior. *Neuroscience Letters, 517*(2), 65–70. https://doi.org/10.1016/j.neulet.2012.02.082

Dion, K., Berscheid, E., & Walster, E. (1972). What is beautiful is good. *Journal of Personality and Social Psychology, 24*(3), 285–290. https://doi.org/10.1037/h0033731

Feingold, A. (1992). Good-looking people are not what we think. *Psychological Bulletin, 111*, 304–341. https://doi.org/10.1037/0033-2909.111.2.304

Frid, A., & Manevitz, L. M. (2020). Analyzing cognitive processes from complex neurophysiologically based data: some lessons. *Annals of Mathematics and Artificial Intelligence, 88*(11), 1125–1153. https://doi.org/10.1007/s10472-019-09669-z

Gerger, G., Leder, H., Tinio, P. P. L., & Schacht, A. (2011). Faces versus patterns: Exploring aesthetic reactions using facial EMG. *Psychology of Aesthetics, Creativity, and the Arts, 5*(3), 241–250. https://doi.org/10.1037/a0024154

Golle, J., Mast, F. W., & Lobmaier, J. S. (2014). Something to smile about: the interrelationship between attractiveness and emotional expression. *Cognition and Emotion, 28*(2), 298–310. https://doi.org/10.1080/02699931.2013.817383

Grootswagers, T., Wardle, S. G., & Carlson, T. A. (2017). Decoding Dynamic Brain Patterns from Evoked Responses: A Tutorial on Multivariate Pattern Analysis Applied to Time Series Neuroimaging Data. *Journal of Cognitive Neuroscience, 29*(4), 677–697. https://doi.org/10.1162/jocn_a_01068

Hajcak, G., MacNamara, A., & Olvet, D. M. (2010). Event-related potentials, emotion, and emotion regulation: an integrative review. *Dev Neuropsychol, 35*(2), 129–155. https://doi.org/10.1080/87565640903526504

Hajcak, G., Weinberg, A., MacNamara, A., & Foti, D. (2011). ERPs and the Study of Emotion. In E. J. Kappenman, & S. J. Luck (Eds.), *The Oxford Handbook of Event-Related Potential Components* (pp. 516–554). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780195374148.013.0222.

Halit, H., de Haan, M., & Johnson, M. H. (2000). Modulation of event-related potentials by prototypical and atypical faces. *Neuroreport, 11*(9), 1871–1875. https://doi.org/10.1097/00001756-200006260-00014

Han, S., Liu, S., Gan, Y., Xu, Q., Xu, P., Luo, Y., & Zhang, L. (2020). Repeated exposure makes attractive faces more attractive: Neural responses in facial attractiveness judgement. *Neuropsychologia, 139*, Article 107365. https://doi.org/10.1016/j.neuropsychologia.2020.107365

Höller, Y. (2021). Quantitative EEG in Cognitive Neuroscience. *Brain Sciences, 11*(4). https://doi.org/10.3390/brainsci11040517

van Hooff, J. C., Crawford, H., & van Vugt, M. (2011). The wandering mind of men: ERP evidence for gender differences in attention bias towards attractive opposite sex faces. *Social Cognitive and Affective Neuroscience, 6*(4), 477–485. https://doi.org/10.1093/scan/nsq066

ISO/CIE 11664–4, 2019, Colorimetry — Part 4: CIE 1976 L*a*b* colour space. ⟨https://www.iso.org/standard/74166.html⟩.

Jackson, L. A., Hunter, J. E., & Hodge, C. N. (1995). Physical attractiveness and intellectual competence: a meta-analytic review. *Social Psychology Quarterly, 58*(2), 108–122. https://doi.org/10.2307/2787149

Johnston, V. S., & Oliver-Rodriguez, J. C. (1997). Facial beauty and the late positive component of event-related potentials. *The Journal of Sex Research, 34*(2), 188–198. https://doi.org/10.1080/00224499709551884

Langlois, J. H., Kalakanis, L., Rubenstein, A. J., Larson, A., Hallam, M., & Smoot, M. (2000). Maxims or myths of beauty? A meta-analytic and theoretical review. *Psychological Bulletin, 126*(3), 390–423. https://doi.org/10.1037/0033-2909.126.3.390

Ma, F. L., Xu, F., & Luo, X. M. (2015). Children's and Adults' Judgments of Facial Trustworthiness: The Relationship to Facial Attractiveness. *Perceptual and Motor Skills, 121*(1), 179–198. https://doi.org/10.2466/27.22.PMS.121c10×1

Ma, F. L., Xu, F., & Luo, X. M. (2016). Children's facial trustworthiness judgments: agreement and relationship with facial attractiveness. *Frontiers in Psychology, 7*, Article 499. https://doi.org/10.3389/fpsyg.2016.00499

Ma, Q., Qian, D., Hu, L. F., & Wang, L. (2017). Hello handsome! Male's facial attractiveness gives rise to female's fairness bias in Ultimatum Game scenarios-An ERP study. *PLoS One, 12*(7), Article e0180459. https://doi.org/10.1371/journal.pone.0180459

Ma, Q., Zhang, L., Pei, G., & Abdeljelil, H. (2017). Neural process of the preference cross-category transfer effect: evidence from an event-related potential study. *Scientific Reports, 7*(1), 3177. https://doi.org/10.1038/s41598-017-02795-w

Ma, Q. G., Hu, Y., Jiang, S. S., & Meng, L. (2015). The undermining effect of facial attractiveness on brain responses to fairness in the Ultimatum Game: an ERP study (Article) *Frontiers in Neuroscience, 9*, Article 77. https://doi.org/10.3389/fnins.2015.00077.

Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. *Journal of Neuroscience Methods, 164*(1), 177–190. https://doi.org/10.1016/j.jneumeth.2007.03.024

Marzi, T., & Viggiano, M. P. (2010). When memory meets beauty: Insights from event-related potentials. *Biological Psychology, 84*(2), 192–205. https://doi.org/10.1016/j.biopsycho.2010.01.013

Munoz, F., & Martin-Loeches, M. (2015). Electrophysiological brain dynamics during the esthetic judgment of human bodies and faces. *Brain Res, 1594*, 154–164. https://doi.org/10.1016/j.brainres.2014.10.061

Oliver-Rodriguez, J. C., Guan, Z. Q., & Johnston, V. S. (1999). Gender differences in late positive components evoked by human faces. *Psychophysiology, 36*(2), 176–185. https://doi.org/10.1142/S0219635212500306

Olson, I. R., & Marshuetz, C. (2005). Facial attractiveness is appraised in a glance. *Emotion, 5*(4), 498–502. https://doi.org/10.1037/1528-3542.5.4.498

Oostenveld, R., Fries, P., Maris, E., & Schoffelen, J. M. (2011). FieldTrip: Open Source Software for Advanced Analysis of MEG, EEG, and Invasive Electrophysiological Data. *Computational Intelligence and Neuroscience, 9*, Article 156869. https://doi.org/10.1155/2011/156869

Oosterhof, N.N., & Todorov, A. (2008). The functional basis of face evaluation. *Proceedings of the National Academy of Sciences, 105* (32), 11087–11092. https://doi.org/doi:10.1073/pnas.0805664105.

Principe, C. P., & Langlois, J. H. (2011). Faces differing in attractiveness elicit corresponding affective responses. Article Pii 927585387 *Cognition & Emotion, 25*(1), 140–148. https://doi.org/10.1080/02699931003612098.

Pronk, T. M., & Denissen, J. J. A. (2020). A rejection mind-set: choice overload in online dating. *Social Psychological and Personality Science, 11*(3), 388–396. https://doi.org/10.1177/1948550619866189

Rossion, B., & Jacques, C. (2008). Does physical interstimulus variance account for early electrophysiological face sensitive responses in the human brain? Ten lessons on the N170. *Neuroimage, 39*(4), 1959–1979. https://doi.org/10.1016/j.neuroimage.2007.10.011

Roye, A., Hofel, L., & Jacobsen, T. (2008). Aesthetics of Faces Behavioral and Electrophysiological Indices of Evaluative and Descriptive Judgment Processes. *Journal of Psychophysiology, 22*(1), 41–57. https://doi.org/10.1027/0269-8803.22.1.41

Rugg, M. D., Furda, J., & Lorist, M. (1988). The effects of task on the modulation of event-related potentials by word repetition. *Psychophysiology, 25*(1), 55–63. https://doi.org/10.1111/j.1469-8986.1988.tb00958.x

Schacht, A., Werheid, K., & Sommer, W. (2008). The appraisal of facial beauty is rapid but not mandatory. *Cognitive Affective & Behavioral Neuroscience, 8*(2), 132–142. https://doi.org/10.3758/cabn.8.2.132

Schein, S. S., & Langlois, J. H. (2015). Unattractive infant faces elicit negative affect from adults. *Infant Behavior & Development, 38*, 130–134. https://doi.org/10.1016/j.infbeh.2014.12.009

Tanaka, H. (2021). Lip color affects ERP components in temporal face perception processing. *JIN, 20*(4), 1029–1038. https://doi.org/10.31083/j.jin2004104

Todorov, A., & Engell, A. D. (2008). The role of the amygdala in implicit evaluation of emotionally neutral faces. *Social Cognitive and Affective Neuroscience, 3*(4), 303–312. https://doi.org/10.1093/scan/nsn033

Treder, M. S. (2020). MVPA-Light: A Classification and Regression Toolbox for Multi-Dimensional Data. *Frontiers in Neuroscience, 14*. https://doi.org/10.3389/fnins.2020.00289

Trujillo, L. T., Jankowitsch, J. M., & Langlois, J. H. (2014). Beauty is in the ease of the beholding: a neurophysiological test of the averageness theory of facial attractiveness. *Cognitive, Affective, & Behavioral Neuroscience, 14*(3), 1061–1076. https://doi.org/10.3758/s13415-013-0230-2

Werheid, K., Schacht, A., & Sommer, W. (2007). Facial attractiveness modulates early and late event-related brain potentials. *Biological Psychology, 76*(1–2), 100–108. https://doi.org/10.1016/j.biopsycho.2007.06.008

Wiese, H., Altmann, C. S., & Schweinberger, S. R. (2014). Effects of attractiveness on face memory separated from distinctiveness: evidence from event-related brain potentials. *Neuropsychologia, 56*, 26–36. https://doi.org/10.1016/j.neuropsychologia.2013.12.023

Yuan, G. J., Liu, G. Y., & Wei, D. T. (2021). Roles of P300 and Late Positive Potential in Initial Romantic Attraction. *Frontiers in Neuroscience, 15*, Article 718847. https://doi.org/10.3389/fnins.2021.718847

Zhang, Y., Kong, F., Chen, H., Jackson, T., Han, L., Meng, J., Yang, Z., Gao, J., & Najam ul Hasan, A. (2011). Identifying cognitive preferences for attractive female faces: An event-related potential experiment using a study-test paradigm. *Journal of Neuroscience Research, 89*(11), 1887–1893. https://doi.org/10.1002/jnr.22724

Zhang, Z., & Deng, Z. (2012). Gender, facial attractiveness, and early and late event-related potential components. *Journal of Integrative Neuroscience, 11*(4), 477–487. https://doi.org/10.1142/S0219635212500306