Psychophysiology

WILEY

PSYCHOPHYSIOLOGY SPR

ORIGINAL ARTICLE OPEN ACCESS

Event-Related Potentials to Facial Expressions Are Related to Stimulus-Level Perceived Arousal and Valence

Amie J. Durston | Roxane J. Itier

Department of Psychology, University of Waterloo, Waterloo, Ontario, Canada

Correspondence: Amie J. Durston (ajdurston@uwaterloo.ca)

Received: 7 November 2024 | **Revised:** 27 January 2025 | **Accepted:** 27 February 2025

Funding: This research was funded by the Natural Sciences and Engineering Research Council of Canada (Grant 418431) and the Canada Foundation for Innovation (Grant 213322) awarded to RJI. AJD was supported by NSERC Graduate Scholarships (Masters and Doctorial) and the Ontario Graduate Scholarship (Masters).

Keywords: arousal | event-related potentials | facial expressions | mass univariate analyses | valence

ABSTRACT

Facial expressions provide critical details about social partners' inner states. We investigated whether event-related potentials (ERP) related to the visual processing of facial expressions are modulated by participants' perceived arousal and valence at the stimulus level. ERPs were recorded while participants ($N=80$) categorized the gender of faces expressing fear, anger, happiness, and no emotion. Participants then viewed each face again and rated them on arousal and valence using 1–9 Likert scales. For each participant, ratings of each unique face were linked back to corresponding ERP trials. ERPs were analyzed at all time points and electrodes using hierarchical mass univariate statistics. Three different ANOVA models were employed: the original emotion model, and models with valence or arousal ratings as trial-level regressors. Results from models with ratings highly overlapped with the original model, although they were more temporally restricted. The N170 component was the most impacted by arousal and valence ratings, with four out of six emotion contrasts revealing significant valence or arousal interactions. Emotion effects on the P2 component were mostly unrelated to ratings. On the EPN component, only two contrasts related to both arousal and valence ratings. Thus, ERP emotion effects are related to participants' perceived arousal and valence of the stimuli, although this association depends on the contrast analyzed. These findings, their limitations, and generalizability are discussed in reference to existing theories and literature.

1 | Introduction

Facial expressions are a critical source of social information, providing details about others' internal states (Keltner et al. 2019). The neural time course of expression processing has been extensively studied using Event-Related Potentials (ERPs) and is often assumed to relate to expressions' perceived valence and arousal (e.g., Calvo, Marrero, et al. 2013; Calvo and Nummenmaa 2016; Durston and Itier 2021; Han et al. 2021; Schupp, Ohman, et al. 2004). However, this assumption has never been directly tested within participants, a gap that the present study addresses.

The first ERP component of interest is the P1, a positive ERP component around 100 ms following face onset on parietal–occipital electrodes. The P1 is sensitive to low-level visual processing (i.e., luminance, pixel intensity, spatial frequency; Rossion and Jacques 2011). P1 modulations by emotion are often interpreted as attentional capture due to low-level stimulus properties (Müller-Bardorff et al. 2018; Palermo and Rhodes 2007; Vuilleumier and Pourtois 2007), although they are highly variable and infrequently replicated (Schindler and Bublatzky 2020).

The N170 is a negative ERP component on lateral posterior electrodes 130–180 ms after face onset, indexing the face structural encoding (Rossion and Jacques 2011). While the N170 is more negative for threat-related (fear, anger) faces than for other expressions (Hinojosa et al. 2015; Schindler and Bublatzky 2020), N170 differences between neutral and other expressions (i.e., happiness) are less clear-cut (Schindler and Bublatzky 2020 for review). Furthermore, few studies have compared angry and fearful faces and find either a more negative N170 for fear (Almeida et al. 2016; Batty and Taylor 2003; Turano et al. 2017) or no difference between the two expressions (Calvo and Beltrán 2013; Herbert et al. 2013; Smith et al. 2013). Although some patterns emerge, it is still inconclusive if, and how, facial expressions modulate the N170.

Following the N170 on posterior electrodes is the visual P2 (~200 ms post face onset). There is limited literature on its role in expression processing, as only 18% of these studies analyze it (Schindler and Bublatzky 2020). Recent work showed that the largest expression effects during early face processing were between the N170 and the P2, where fearful and angry faces elicited more negative amplitudes compared to happy and neutral faces (Durston and Itier 2021; Han et al. 2021; Itier and Durston 2023; Qiu et al. 2023; but see Hudson et al. 2021).

Lastly, the early posterior negativity (EPN), which follows the P2, indexes the allocation of attentional resources toward emotional and threatening stimuli (Ohman 2009; Schupp, Ohman, et al. 2004). When measured as a mean amplitude across a predefined time window (typically 200–350 ms), the EPN is more negative for fearful and angry faces than for other expressions (Schindler and Bublatzky 2020), and the few studies that have compared fearful and angry expressions have reported null EPN results (Herbert et al. 2013; Smith et al. 2013). Recent mass univariate analyses revealed EPN differences between emotional and neutral expressions (Itier and Durston 2023), but no difference between emotional categories (Durston and Itier 2021; Hudson et al. 2021; Itier and Durston 2023).

The above-mentioned emotion effects are often interpreted through the lens of affective content. The perceived intensity of facial expressions, which is often discussed as reflecting arousal, may increase P1 amplitude (Müller-Bardorff et al. 2018). Many have proposed that N170 emotion effects are related to the discrimination of valence (Batty and Taylor 2003; Calvo and Beltrán 2013) or arousal (Calvo and Nummenmaa 2016; Williams et al. 2006). Likewise, studies have speculated that the P2 indexes extraction of affective content (e.g., arousal and valence) from faces and words (Calvo, Marrero, et al. 2013; Durston and Itier 2021; Han et al. 2021) and that EPN emotion results are linked to arousal (Calvo and Beltrán 2013; Langeslag and van Strien 2018; Schupp, Ohman, et al. 2004) or valence discrimination (Calvo and Nummenmaa 2016). However, only two studies have directly tested these interpretations. Almeida et al. (2016) manipulated the arousal content of stimuli across multiple facial expressions and found that as stimuli arousal increased, so did the N170 amplitude. Li et al. (2022) took a computational decoding approach across the entire time course and found that amplitudes in the N170 time window were related to the perceived arousal of the stimuli, while amplitudes in the EPN time window related to both valence and arousal, with no affective content extraction during the P2 time window. Critically, both studies (Almeida et al. 2016; Li et al. 2022) used valence and arousal ratings from different participants than their ERP participants. Thus, no study to date has directly related perceived arousal or valence of facial expressions to ERP amplitudes in the same participants, a critical gap we addressed in the present study.

Furthermore, many inconsistencies remain across the emotion processing field, possibly due to methodological factors. First, most ERP studies on the topic are underpowered, with low numbers of trials (< 50 trials per condition) and small sample sizes (often < 30 participants). Second, the statistics used by most of these studies are error-prone (Luck and Gaspelin 2017), limiting the reliability of the findings. Third, studies typically use a small number of faces repeated a large number of times, with different face databases possibly eliciting different behavioral or neural responses due to different perceived emotional intensities of the faces (Adolph and Alpers 2010). Lastly, most experiments do not control gaze location, despite the fact that it alters the neural processing of faces and facial expressions (Itier and Durston 2023; Siklos-Whillans and Itier 2024).

The current study used a within-subjects design to directly compare the neural processing of angry, fearful, happy, and neutral faces using ERPs and investigate whether expression differences on ERPs were related to participants' perceived valence and arousal of these expressions. We controlled gaze location using an eye tracker, collected a large sample ($N = 80$), included a large number of faces from four databases, and used robust Mass Univariate (MU) statistics. Based on the above literature, we predicted that: (1) more negative N170, P2, and EPN amplitudes would be found for fearful and angry expressions compared to happy and neutral faces; (2) N170-P2 effects would be related to perceived arousal; (3) EPN effects would be linked to both perceived arousal and valence.

2 | Methods

The study design and hypotheses were preregistered on the Open Science Framework (OSF) and deviations from the preregistration are disclosed. Final analyses output and detailed analyses steps are also available on OSF (https://osf.io/zyh3e/). The individual participant data collected for this study are available from the corresponding author upon reasonable request.

2.1 | Participants

Ninety-five participants were recruited through the University of Waterloo (UW) between May and December 2023. Participants were remunerated with cash (up to $CAN36) or course credit and cash (up to $CAN12). All participants had normal or corrected-to-normal vision and reported no history of psychiatric or neurological disorders, brain injury, or lesions. They had been in North America for more than 10 years and were thus used to Caucasian faces and expressions. Following the Declaration of Helsinki, all participants signed an informed consent letter before the study, which was approved by the UW Ethics Board (project #41702). Two participants were rejected for falling asleep, 11 were rejected due to corrupted eye-tracking data, and two were rejected due to quitting following the first task (see procedures). This left a total of 80 participants in the present analyses (Table 1). No preliminary sample size calculations were computed, but others have deemed 80 participants adequate (e.g., Schindler et al. 2023).

2.2 | Stimuli

Forty Caucasian identities were used, comprised of 10 (5 male) from four databases: NIMSTIM (Tottenham et al. 2009), Chicago Face Database (Ma et al. 2015), Radboud Face Database (Langner et al. 2010), and FACES Database (Ebner et al. 2010). Only faces with minimal makeup or facial hair and validated high emotion recognition accuracy were used. Fearful, happy, angry, and neutral expressions were used from each of the chosen identities. Fearful and happy expressions had open mouths, and angry and neutral faces had closed mouths.

Faces were edited with GIMP v2.10.32. Colored stimuli were cropped into an ellipse with its middle sitting just above the tip of the nose and were then centered on a white background. Any potentially attention-grabbing features were edited out. Colored images were equated for contrast and pixel intensity[1] using the SHINE toolbox (Willenbockel et al. 2010). Photographs were resized so faces sustained 5.39°w × 7.71°h of visual angle when sitting 70 cm from the screen.

2.3 | Procedure

Participants were fitted with an EEG cap and completed demographics and additional questionnaires.[2] Participants were seated 70 cm from the computer screen, with their heads on a chin rest. Prior to the experiment, participants completed baseline brain recordings (not reported here). The first task was gender discrimination, orthogonal to the emotion variable of interest. Participants fixated on a centered cross for 200 ms (verified by an eye tracker), then a face was presented for 250 ms, followed by a 2000 ms fixation cross screen allowing responses using arrow keys (counterbalanced across participants). On each trial, the center of the face's nose was situated where the fixation cross was. Eight practice trials were given before the experimental trials. The experiment consisted of 12 blocks of 40 faces, for a total of 480 trials (120/condition). Each block consisted of stimuli from the same database (10 identities × 4 expressions) and was repeated three times (4 database blocks × 3 repetitions). Databases were separated by block to ensure block homogeneity. Blocks and faces within blocks were presented in a randomized order. Eye tracking was recalibrated after six blocks or fewer if needed.

After the gender task, participants completed a rating task, in which no EEG was recorded. Participants were given verbal and written instructions to rate the perceived valence and arousal of the face. Participants viewed each unique stimulus again, using the same timing and trial progression as the gender task. They had 10 s after the face disappeared to rate the face on how positive/negative it was, and how aroused the face was on scales from 1 to 9 using the computer mouse (Figure 1). These prompts were chosen based on previous studies (e.g., Calvo, Marrero, et al. 2013). Participants completed four blocks of 40 faces, for

TABLE 1 | Demographic details from the current sample ($N = 80$).

	Biological sex		Dominant hand		Racial background			
Mean age (SD)	**# Males (%)**	**# Females (%)**	**# Right-handed (%)**	**# Left-handed (%)**	**# East Asian (%)**	**# Caucasian (%)**	**# South Asian (%)**	**# Other[a] (%)**
21 (3.0)	40 (50%)	40 (50%)	71 (88.8%)	9 (11.2%)	26 (32.5%)	22 (27.5%)	18 (22.5%)	14 (17.5%)

[a]Other racial groups include 4 (5.0%) middle eastern, 2 (2.5%) black/African, and 8 (10%) mixed race participants.

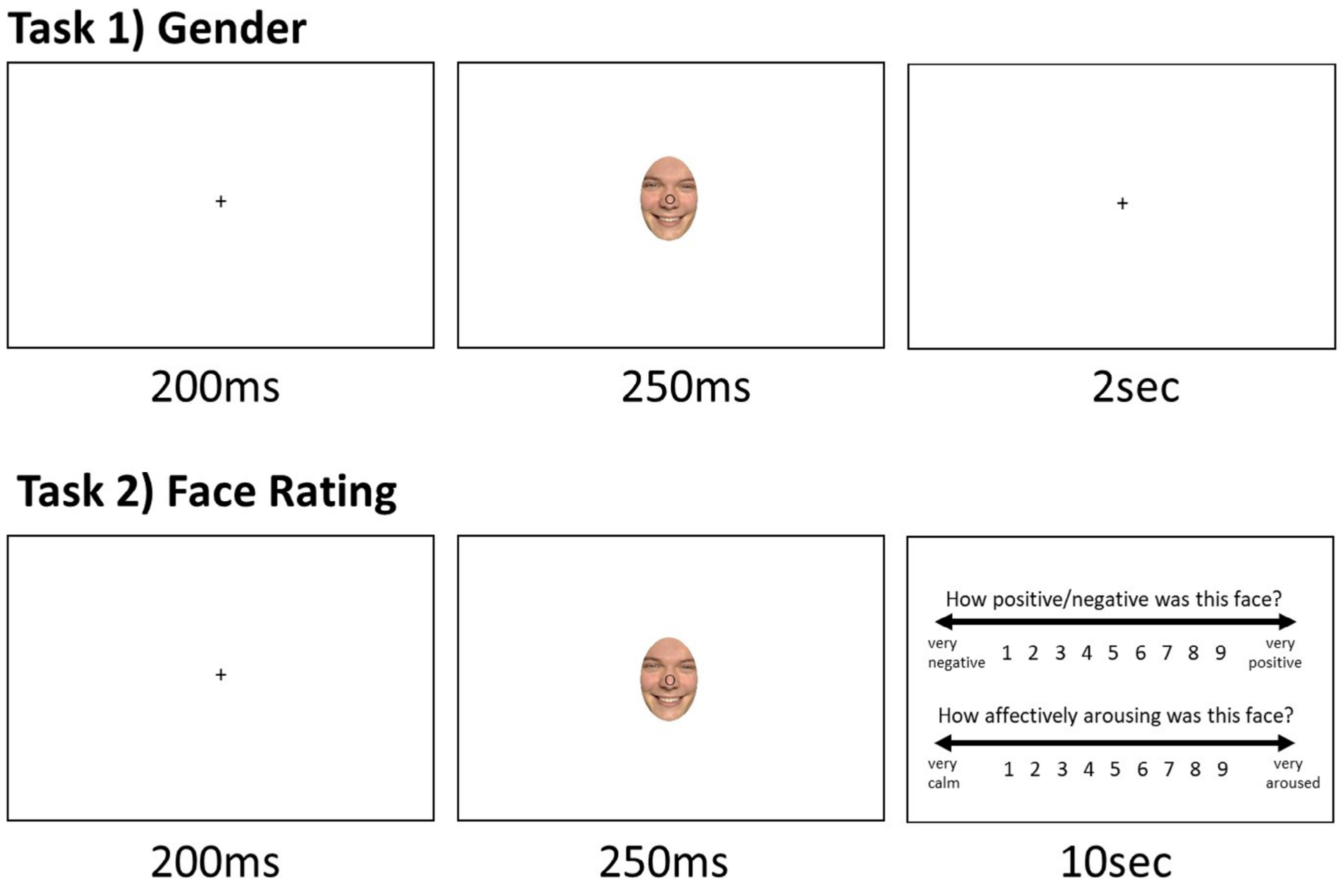

FIGURE 1 | Tasks and trial progression. The black circle on the face represents the 1.4° region of interest participants had to fixate within for the trial to be kept during analysis. The circle itself was not visible on the faces during the experiment. Trials where participants saccaded outside of this region during face presentation were rejected (see Methods). The face presented in this figure is identity 066 from the FACES database, which is approved for use in academic publications (Ebner et al. 2010).

a total of 160 trials. The order of presentation (which rating was presented on the top of the screen) was the same from trial to trial and was counterbalanced across participants. Eye tracking was recalibrated between each task, and after any block with complications. Then, participants completed a personal affective rating (how each face made them feel) and an emotion categorization task (results for these two additional tasks are not reported here).

2.4 | Eye Tracking and EEG Acquisition

Eye movements were monitored using an EyeLink 1000 remote eye tracker (SR Research), sampling at 1000 Hz. Calibration was done using the dominant eye and a 9-point automated calibration test. A drift correction was used if participants took longer than 10 s to fixate on the fixation cross at the start of the trial, and a mid-block recalibration was performed after multiple drift corrections.

EEG was continuously recorded during the gender task at 500 Hz by an Acti64Champ system (Brain Vision Solutions Inc.). Custom caps were used, following the 10–20 extended system, including 64 electrode channels. Electrode sites F1, F2, AF3, and AF4 were excluded from the recording montage, while PO9/PO10 and TP9/TP10 were added. Electrode Cz was used as the online reference, but data were re-referenced offline to the common average reference.

2.5 | Data Processing

EEG trials were rejected (1) if a saccade exceeding the 1.4° region of interest centered on the face's nose was recorded (Figure 1) and (2) if participants did not complete either task (gender and rating) for the corresponding face within the allotted time (2 and 10 s, respectively). Overall, this led to the rejection of 22.28 (SD = 17.96) trials per condition on average.[3] We preregistered rejecting trials with incorrect gender discrimination and emotion categorization responses; however, for the sake of power, we included all trials. Recent work has shown that attentional task demands do not impact the early neural processing of expressions (Durston and Itier 2021; Hudson et al. 2021; Itier and Durston 2023; Itier and Neath-Tavares 2017), so we concluded that rejection based on orthogonal (gender) and explicit (emotion) task accuracy was unnecessary.

Remaining trials were processed using MATLAB toolboxes EEGLab (Delorme and Makeig 2004) and ERPLab (Lopez-Calderon and Luck 2014) with −100 ms to +350 ms epochs around face onset. Data were filtered using a 0.01–30 Hz bandpass filter.

Trials with artifacts exceeding ±70 μV were rejected automatically. Participants who had a high number of trials rejected were inspected for noisy channels, which were interpolated as needed. The remaining trials were visually inspected and rejected when artifacts were visible. These cleaning steps resulted in an additional 8.64 (SD = 12.35) trials per condition rejected on average. The final cleaned datasets for the 80 participants had an average of 86.30 (SD = 20.68) trials per condition.[4]

2.6 | Data Analysis

Behavioral data analyses were conducted in JASP (version 0.17.2.1). Each participant's average valence and arousal rating for each expression were entered into traditional and Bayesian repeated measures ANOVAs. The Greenhouse–Geisser correction was used when sphericity was violated. We set up the Bayesian model with the default expected priors favoring the alternative hypothesis (i.e., BF_{10}). BF_{10} values above 1 provide evidence for the alternative hypothesis, with values above 100 providing extreme evidence (Lee and Wagenmakers 2014). Average arousal ratings were normally distributed in terms of skew (< |3|) and kurtosis (< |10|), as were average valence ratings, except for neutral expressions (kurtosis = 11) which had slightly over-concentrated ratings around the midline of the distribution (Kline 1998). However, this shows that the faces were perceived as truly neutral in valence. Given the trial-by-trial nature of the ERP analyses (see below) and the small deviation from the recommended guidelines, no transformations were conducted prior to analyses.

ERP data were analyzed using LIMO EEG (EEG-Master version; Pernet et al. 2011). This toolbox takes a hierarchical approach, where the first level computes a regression model based on each participant's trial-level data, at each time point and electrode. This step outputs full scalp (all electrodes × time points) matrices of regression outputs (i.e., estimated beta parameters), which are then entered into the group level statistical model. This method ensures participants' trial-level variance is accounted for. For more details on the fundamentals behind this toolbox, see Pernet et al. (2011).

Three different regression models were computed at the first level using weighted least squares. The "original" model included only the expression categories, revealing the electrodes and time points at which the expressions significantly differed from each other. The "arousal" model included the emotion categories and arousal ratings as a regressor. That is, for each participant, the ratings obtained for each face during the rating task were linked to the appropriate trials during the EEG recordings of the gender task (i.e., participant 1 rating of face 1 was used as a regressor for each trial where face 1 was presented to that participant). Subsequently, the first level "arousal" model outputs each participant's beta parameters associated with the interaction between arousal ratings and emotional expression on EEG activity. Lastly, a "valence" model was computed in the same manner as the arousal model, but with valence ratings as the regressor.[5]

At the second level, electrode locations were defined using a customized neighborhood matrix (electrode distance = 0.4799). Outputs from the three models first levels were then each entered into a data-driven, whole-epoch mass univariate ANOVA (1 × 4 emotions), with alpha set to 0.05, 1000 bootstraps, and a cluster mass correction (Maris and Oostenveld 2007; Pernet et al. 2015). For each model, follow-up *F*-contrasts (comparing two expressions) were run on the whole epoch, using Bonferroni-corrected *p*-values of 0.008 to account for the six contrasts conducted. All tests were run from −100 to +350 ms at all electrodes (225 data points × 64 electrode). Note that Hotelling *t*-tests are used in these models to account for the covariance between measures; thus, there is no need to adjust for sphericity in the ERP data.

Second-level results from the original model indicate traditional ANOVA emotion effects on ERP amplitudes. The arousal and valence models, including ratings at the stimulus level, indicate whether the relationship between ratings and EEG data differs between emotions. Conceptually, this represents an interaction between emotion category and respective ratings on ERP amplitudes, allowing us to infer which emotion effects from the original model are related to participants' perceived arousal or valence of each face (see Pernet et al. 2021 for a similar model). Note that these hierarchical interaction models were used in lieu of the preregistered emotion × ratings ANCOVA model to preserve the stimuli-level nature of the dataset. Additional valence and arousal ERP correlation analyses were preregistered, and results are included in the Figure S2.

ERPs and difference wave graphs were created in LIMO EEG. Chosen measures of central tendency were standard means within subjects, and 20% trimmed means across subjects. Confidence intervals were computed with a Bonferroni corrected alpha adjusted for the number of follow-up tests ($\alpha = 0.008$). Furthermore, LIMO uses Bayesian Highest Density Interval (HDI) as a confidence interval around the difference wave, which allows us to be confident in the null when plotting the results if the difference wave and its confidence interval include zero. Electrodes where *F*-values for statistical clusters were maximal were plotted in the present figures to maintain a data-driven approach. Lastly, effect sizes were computed using limo_get_effect_size.m (Pernet et al. 2011), which outputs an effect size tailored for the analysis type, in this case the multivariate and boundless Mahalanobis Distance (*D*).

3 | Results

3.1 | Behavioral Ratings

Participants rated the facial expressions as expected. There was a main effect of emotion category on mean arousal ratings ($F(2.158, 170.451) = 198.498$, $p < 0.001$, $\eta_p^2 = 0.715$, $BF_{10} = 2.512 \times 10^{66}$, $R^2 = 0.683$). Post hoc tests showed that all emotions ratings were significantly different from each other, except for angry and happy faces ($p = 0.597$, $BF_{10} = 0.318$; see Figure S1). Fearful faces were rated highest in arousal, followed by angry, happy, and, lastly, neutral faces (Table 2). There was also a main effect of emotion category on valence ratings ($F(1.774, 140.179) = 939.104$, $p < 0.001$, $\eta_p^2 = 0.922$, $BF_{10} = 9.221 \times 10^{154}$, $R^2 = 0.902$), with all expressions differing significantly from each other (Figure S1). Happy faces were rated the most positive, followed by neutral, fearful, and, lastly, angry faces (Table 2).

3.2 | ERP Mass Univariate ANOVA Analyses

3.2.1 | Main Effects

All three models revealed a significant effect of emotion and had a large amount of spatial and temporal overlaps. All models included the P1, N170, P2, and EPN components, being significant from 74 to 348 ms and on a majority of the scalp (Figure 2; Table 3). Interpretations cannot be made based on main effects alone; as with all ANOVA models, follow-up contrasts explain the story much better, and this is what we turn to next.

TABLE 2 | Means, standard deviations, minimum and maximum values for arousal, and valence ratings of each expression.

	Arousal				Valence			
	Mean	**SD**	**Min**	**Max**	**Mean**	**SD**	**Min**	**Max**
Angry	5.79[a]	1.11	2.81	7.96	2.9	0.56	1.66	4.68
Fear	6.63	0.99	3.75	8.52	3.31	0.58	1.78	4.80
Happy	5.51[a]	1.45	1.17	8.52	7.26	0.71	5.32	8.91
Neutral	2.70	1.32	1.00	5.31	4.62	0.49	2.05	5.97

[a]Means marked with do not significantly differ from each other.

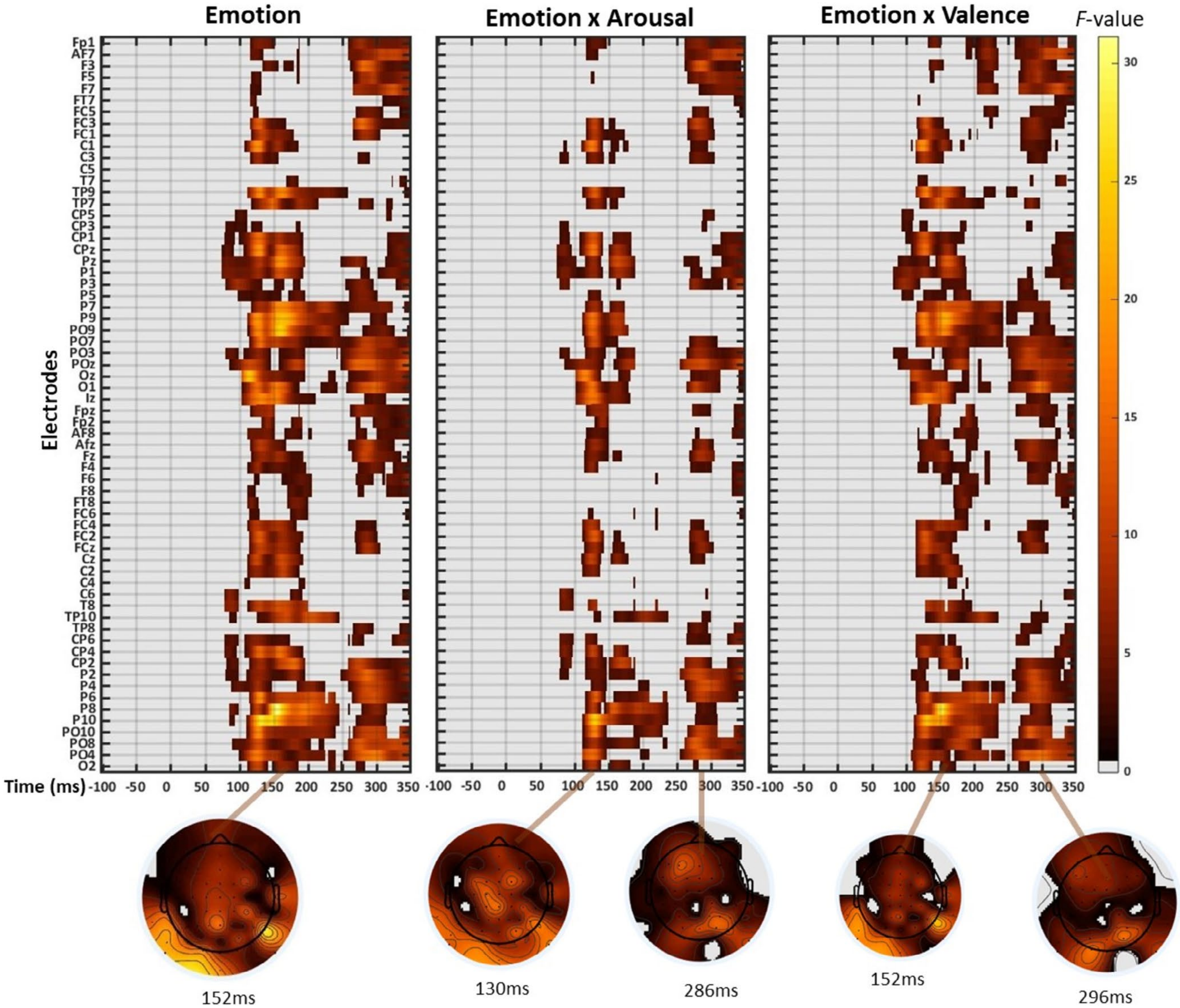

FIGURE 2 | Main effects of emotion as revealed by original (emotion), arousal, and valence models ($p < 0.05$). Time in milliseconds is along the *x*-axes, and electrodes are along the *y*-axes. The color bar represents the *F*-values of significant datapoints, where gray data points are insignificant. Uncorrected (i.e., not filtered using cluster mass) topographic plots underneath the raster plots represent the scalp distribution of the *F*-values at key time points where significant clusters were maximal. Topo plots use the same color scale as the rasters.

TABLE 3 | Main effects summary.

Model	Timing	Cluster max	$F(3, 79)$	p	Mahalanobis distance (D)
Original	74–348 ms	152 ms on P8	31.13	0.001	9.48×10^{-4}
Arousal	74–236 ms	130 ms on P10	25.20	0.001	7.67×10^{-4}
	254–348 ms	286 ms on PO4	13.79	0.001	4.20×10^{-4}
Valence	82–244 ms	152 ms on P8	28.62	0.001	8.71×10^{-4}
	248–348 ms	296 ms on O1	16.11	0.001	4.90×10^{-4}

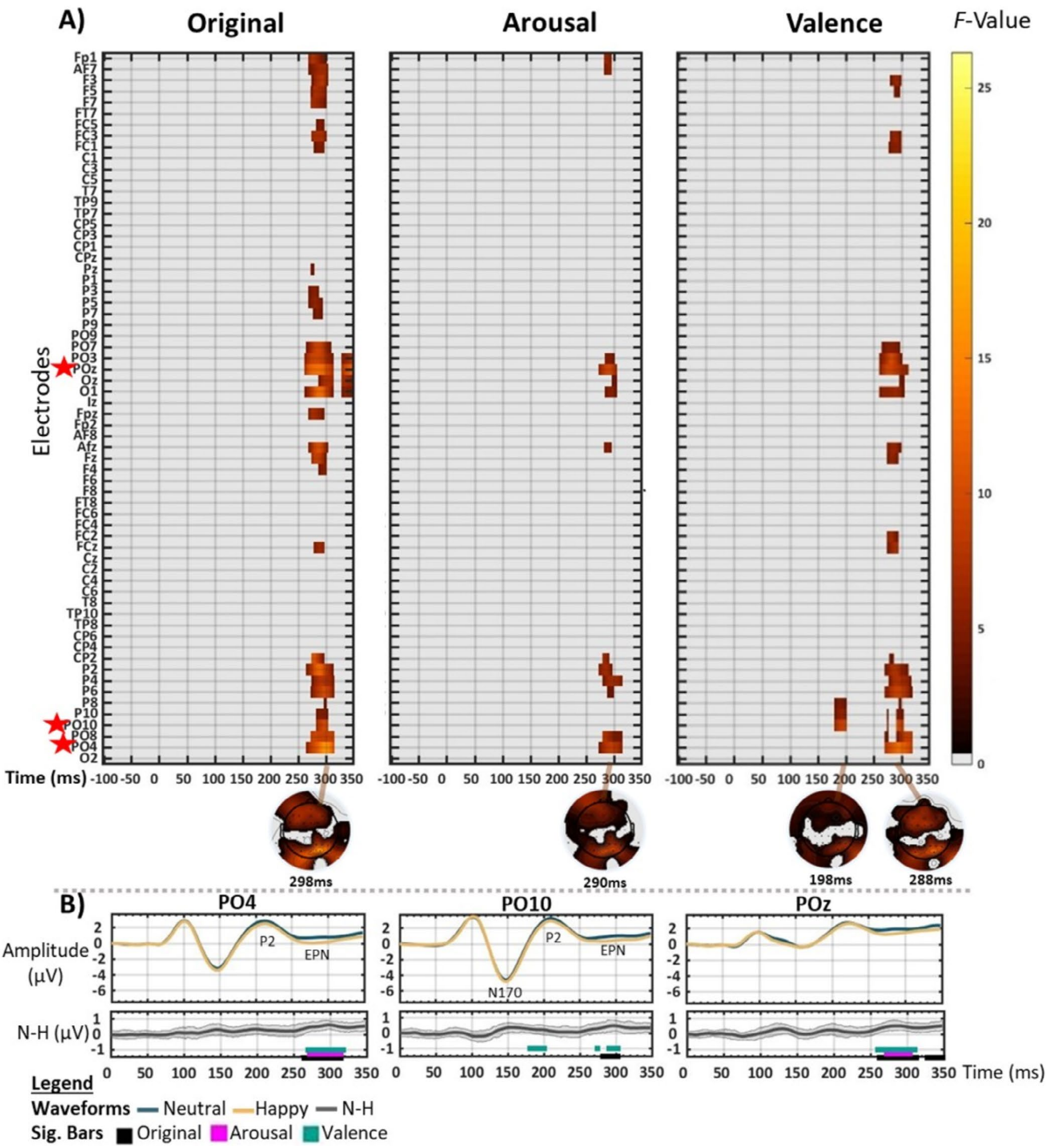

FIGURE 3 | Legend on next page

FIGURE 3 | Results from the neutral and happy expression contrasts ($p < 0.008$). (A) Time in milliseconds is along the x-axes, and electrodes are along the y-axes. The color bar represents the F-values of significant datapoints where gray indicates non-significant datapoints. F-values were corrected using cluster mass correction. Red stars indicate electrodes plotted in B. Uncorrected (i.e., not filtered using cluster mass) topographic plots underneath the raster plots represent the scalp distribution of the F-values at key time points where significant clusters were maximal. Topo plots use the same color scale as the raster. (B) ERPs (top) and difference wave (bottom; neutral-happy) on key electrodes. On the difference wave plots, significance bars indicate when the neutral-happy contrast was significant in the original model (black), in the arousal model (magenta) and in the valence model (cyan). Confidence intervals on the difference wave plots were computed with 20% trimmed means and Bayesian bootstrapping at $\alpha = 0.008$.

TABLE 4 | Neutral versus happy F-contrast summary.

Model	Timing	Electrodes	Cluster max	$F(1, 79)$	p	N−H $\Delta\mu V$	D
Original	262–314 ms	Full scalp	298 ms on PO4	16.65	0.001	0.61	1.03
	328–348 ms	Left PO	348 ms on POz	12.29	0.001	0.53	0.76
Arousal	272–314 ms	Parietal & PO	290 ms on POz	9.87	0.001	0.53	0.61
	272–314 ms	Left frontal	284 ms on Fp1	6.42	0.002	−0.44	0.40
Valence	180–200 ms	Right PO	198 ms on PO10	9.50	0.001	0.24	0.59
	260–318 ms	Posterior	296 ms at PO4	13.23	0.001	0.61	0.82
		Frontal	288 ms on FCz	7.18	0.001	−0.40	0.44

Note: Details of each cluster can be found in Table S1.
Abbreviations: N−H $\Delta\mu V$, neutral minus happy faces amplitude difference; PO, parieto-occipital sites.

3.2.2 | Neutral-Happy Expression Contrasts

In all three models, neutral and happy expressions differed significantly during the EPN timing (250–350 ms; Figure 3; Table 4). Happy expressions elicited more negative amplitudes than neutral expressions on the EPN, an effect related to both arousal and valence ratings. Furthermore, there was a unique valence effect between the N170 and P2 components.

3.2.3 | Neutral-Angry Expression Contrasts

Compared with neutral expressions, angry expressions elicited more negative amplitudes during the P1–N170 interval but less negative amplitudes on the N170–P2 interval, including a more positive P2 component (Figure 4; Table 5). Differences between neutral and angry faces before and during the N170 were related to valence ratings, while arousal was only related to this contrast briefly in the left hemisphere just after the N170 and at centroparietal sites.

3.2.4 | Neutral–Fearful Expression Contrasts

Fearful expressions elicited more negative N170 amplitudes than neutral expressions from 120 to 182 ms (N170-P2 interval, Figure 5; Table 6; Table S3), and this was related to perceived valence, not arousal. The arousal model did not return any significant points.

3.2.5 | Happy–Angry Expression Contrasts

Happy faces elicited more negative amplitudes than angry faces before the N170, but more positive amplitudes than angry faces on the P2 (Figure 6; Table 7; Table S4). Arousal ratings related to happy–angry differences before the N170 and briefly later (164–174 ms) on midline parietal electrodes. There was no influence of valence ratings on these results.

3.2.6 | Happy–Fearful Expression Contrasts

Overall, happy and fearful faces predominantly differed on the EPN, with more negative amplitudes for happy than for fearful faces (Figure 7; Table 8; Table S5). This effect was related to both arousal and valence ratings.

3.2.7 | Angry–Fearful Expression Contrasts

Angry and fearful faces revealed highly overlapping results for all three models (Figure 8; Table 9; for more details, see Table S6). All models were significant during the P1–N170 interval, with more negative amplitudes for fear than for anger, and during the N170–P2 interval, with more negative amplitudes for anger than for fear. These timings were related to both arousal and valence ratings.

4 | Discussion

The present study investigated the neural processing of neutral, happy, angry, and fearful expressions, and whether ERP differences between these expressions were related to their perceived valence or arousal. Using a large sample ($N = 80$), we extended previous research by: (1) conducting a data-driven full-scalp Mass Univariate (MU) analysis; (2) using a gaze-contingent presentation to ensure fixation on the nose; (3) obtaining ratings of face arousal and valence from the same participants whose ERP data were collected. We review our findings below and discuss their implications for understanding face expression processing.

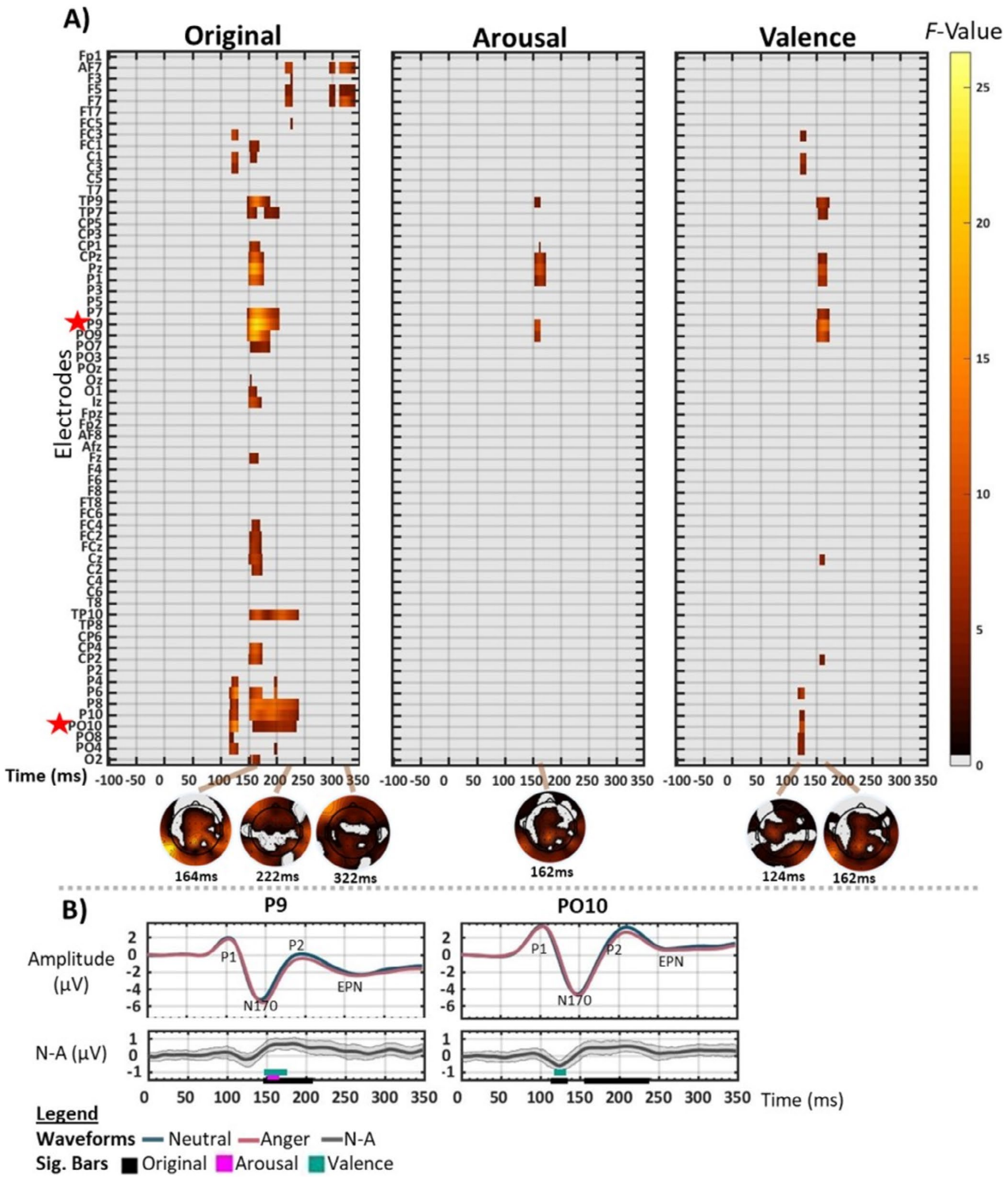

FIGURE 4 | Results from the neutral and angry expression contrasts ($p < 0.008$). (A) Time in milliseconds is along the *x*-axes and electrodes are along the *y*-axes. The color bar represents the *F*-values of significant datapoints, where gray indicates nonsignificant datapoints. *F*-values are corrected using cluster mass correction for multiple comparisons. Red stars beside electrode names indicate electrodes plotted in B. Uncorrected (i.e., not filtered using cluster mass) topographic plots underneath the raster plots represent the scalp distribution of the *F*-values at key time points where significant clusters were maximal. Topo plots use the same color scale as the rasters. (B) ERPs (top) and difference wave (bottom; neutral-angry) on key electrodes. On the difference wave plots, significance bars indicate when the contrast was significant in the original model (black), in the arousal model (magenta) and in the valence model (cyan). Confidence intervals on the difference wave plots were computed with 20% trimmed means and Bayesian bootstrapping at $\alpha = 0.008$.

Our behavioral results are aligned with past work for both valence ratings (Calvo et al. 2016; Calvo, Marrero, et al. 2013; but see Herbert et al. 2013) and arousal ratings, specifically the equivalent arousal ratings for happy and angry faces (Calvo, Marrero, et al. 2013; Herbert et al. 2013). In addition, we report the novel finding that fearful faces were perceived as more aroused than other expressions, while others have reported equal or lower arousal ratings for fearful compared to angry and happy faces (Calvo et al. 2016; Herbert et al. 2013). Since we used near-identical valence and arousal prompts as others (e.g., Calvo et al. 2016), this new result is unlikely due to the procedure but rather likely due to our sample size being roughly four times larger than previous studies ($N=80$ compared to $N_{average}=22$[6]).

4.1 | No Emotion, Arousal or Valence Effect on P1

No effects were found on the visual P1 component itself. The earliest time points were found only for contrasts including angry faces, with less negative amplitudes for angry than for other faces on the downslope between the P1 and N170, that is, after the P1 peak. These effects were closer to the N170 peak than to the P1 and are discussed in more detail in the next section on the N170 component.

Some have found no effects of expression on the P1 (e.g., Calvo, Marrero, et al. 2013; Calvo and Beltrán 2013; Durston and Itier 2021), while others have seen P1 modulations by emotions (Batty and Taylor 2003; Itier and Durston 2023; Palermo and Rhodes 2007; Vuilleumier and Pourtois 2007). Overall, P1 emotion effects are highly inconsistent and are impacted by uncontrolled low-level factors such as pixel intensity and contrast (Rossion and Jacques 2011), but also by attention and the paradigm used (Schindler and Bublatzky 2020). In the present study, neutral, happy, angry, and fearful faces were all comparable on global luminance and pixel intensity. Furthermore, we used a large number of faces from four different face databases to account for potential differences in emotion intensities between databases (Adolph and Alpers 2010). Thus, with the present large sample, well-controlled stimuli, and full-scalp robust analysis, our findings suggest there is no influence of emotion, nor perceived valence or arousal, on the P1.

Exploratory analyses suggested there may be affective influences earlier than the P1 where a correlation existed between happy and fearful expression amplitudes and average arousal ratings from 50 to 109 ms (Figure S2), effects likely attributable to local low-level properties (e.g., bright smile for happy faces, wide white sclera for fearful faces), which were not controlled for. Some participants may be more sensitive to these low-level properties than others, which would cascade along the visual processing pathway all the way down to behavioral perception of arousal. However, these exploratory analyses do not encompass the trial and stimuli-level dynamics of the overall analyses, so these results must be interpreted with caution. Again, when ratings at the stimulus level are taken into account, no early or P1-related modulations by valence or arousal are seen.

TABLE 5 | Neutral versus angry *F*-contrast summary.

Model	Timing	Electrodes	Cluster max	$F(1, 79)$	p	N−A $\Delta\mu V$	D
Original	116–130 ms	Right P	126 ms on PO10	15.88	0.001	−0.60	5.63×10^{-5}
		Left C	124 ms on FC3	9.08	0.005	0.30	5.41×10^{-5}
	148–238 ms	C, T, PO	164 ms on P9	22.41	0.001	0.69	0.16
		Left F	222 ms on AF7	8.84	0.003	−0.48	0.08
Arousal	154–172 ms	Left T, P, & O	160 ms on P9	9.69	0.004	0.67	0.60
		Left and midline P	162 ms on Pz	10.23	0.002	−0.53	0.63
Valence	118–130 ms	Right PO	124 ms on PO10	9.33	0.001	−0.60	0.58
		Left C	124 ms on C1	7.15	0.003	0.36	0.44
	150–172 ms	Central P	160 ms on Pz	9.55	0.001	−0.53	0.59
		Left T and P	162 ms on P9	13.45	0.001	0.68	0.83

Note: Details of each cluster can be found in Table S2.
Abbreviations: C, central sites; F, frontal sites; N−A $\Delta\mu V$, neutral minus angry faces amplitude difference; O, occipital sites; P, parietal sites; PO, parieto-occipital sites; T, temporal sites.

FIGURE 5 | Results from the neutral and fearful expression contrasts ($p<0.008$). (A) Time in milliseconds is along the *x*-axes and electrodes are along the *y*-axes. The color bar represents the *F*-values of significant datapoints, where gray indicates nonsignificant datapoints. *F*-values are corrected using cluster mass correction for multiple comparisons. Red stars beside electrode names indicate electrodes plotted in B. Uncorrected (i.e., not filtered using cluster mass) topographic plots underneath the raster plots represent the scalp distribution of the *F*-values at key time points where significant clusters were maximal. Topo plots use the same color scale as the rasters. (B) ERPs (top) and difference wave (bottom; neutral–fearful) on key electrodes. On the difference wave plots, significance bars indicate when the contrast was significant in the original model (black), in the arousal model (magenta) and in the valence model (cyan). Confidence intervals on the difference wave plots were computed with 20% trimmed means and Bayesian bootstrapping at $\alpha=0.008$.

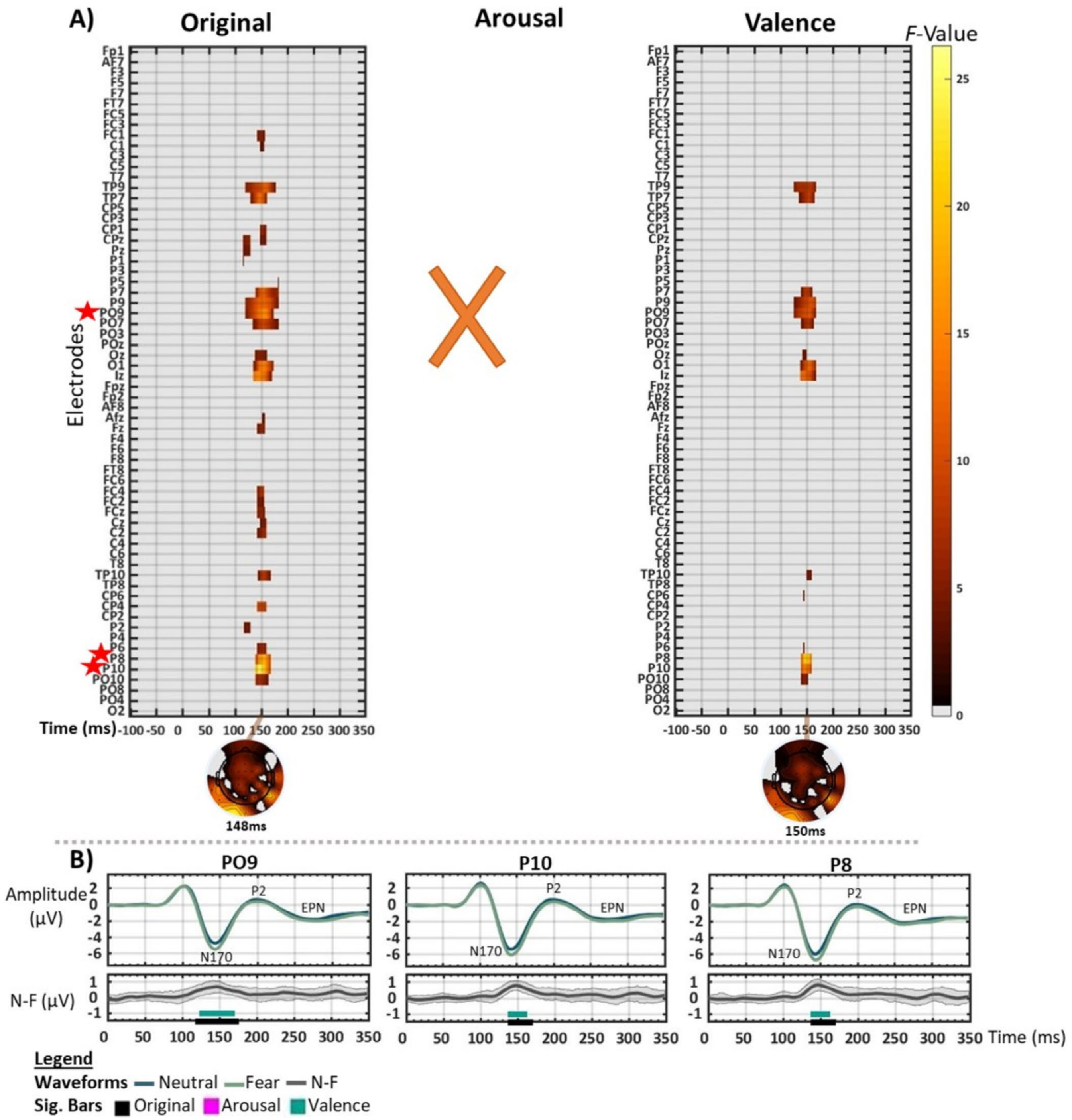

FIGURE 5 | Legend on previous page

TABLE 6 | Neutral versus fear *F*-contrast summary.

Model	Timing	Electrodes	Cluster max	*F*(1, 79)	*p*	N−F $\Delta\mu V$	*D*
Original	120–182 ms	T, P, O	148 ms on P10	25.08	0.001	0.79	1.55
Arousal	N/A						
Valence	126–166 ms	T, P, O	P8 at 150 ms	21.89	0.001	0.82	1.35

Note: Details of each cluster can be found in Table S3.
Abbreviations: N−F $\Delta\mu V$, neutral minus fearful faces amplitude difference; O, occipital sites; P, parietal sites; T, temporal sites.

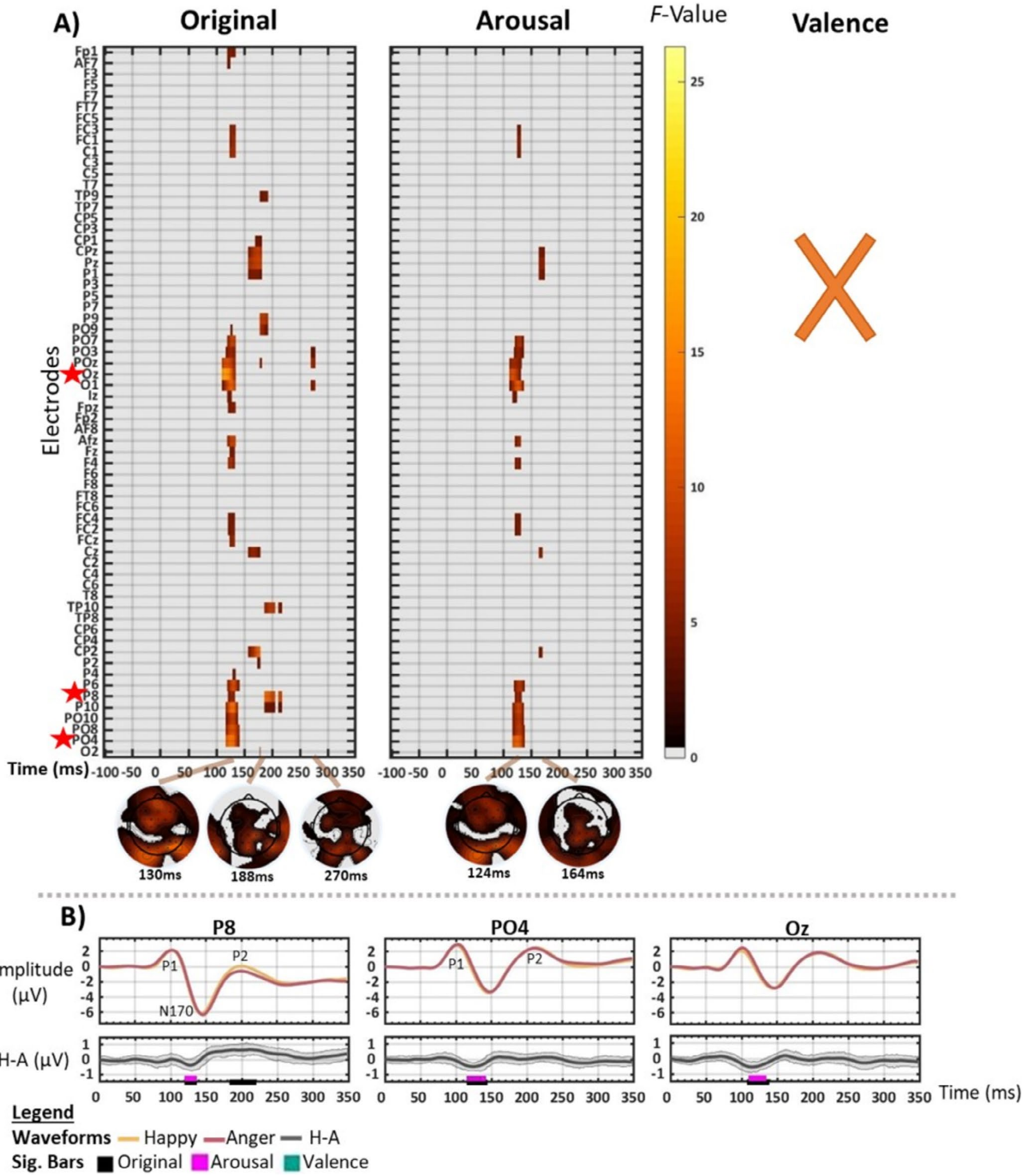

FIGURE 6 | Results from the happy and angry expression contrasts ($p < 0.008$). (A) Time in milliseconds is along the x-axes and electrodes are along the y-axes. The color bar represents the F-values of significant datapoints, where gray indicates non-significant datapoints. F-values are corrected using cluster mass correction for multiple comparisons. Red stars beside electrode names indicate electrodes plotted in B. Uncorrected (i.e., not filtered using cluster mass) topographic plots underneath the raster plots represent the scalp distribution of the F-values at key time points where significant clusters were maximal. Topo plots use the same color scale as the rasters. (B) ERPs (top) and difference wave (bottom; happy–angry) on key electrodes. On the difference wave plots, significance bars indicate when the contrast was significant in the original model (black), in the arousal model (magenta) and in the valence model (cyan). Confidence intervals on the difference wave plots were computed with 20% trimmed means and Bayesian bootstrapping at $\alpha = 0.008$.

TABLE 7 | Happy versus angry F-contrast summary.

Model	Timing	Electrodes	Cluster max	$F(1, 79)$	p	H–A $\Delta\mu V$	D
Original	110–140 ms	P&O	116 ms on Oz	17.33	0.001	−0.54	1.07
	120–134 ms	Frontal	130 ms on AFz	9.42	0.001	0.48	0.58
	158–216 ms	Midline C & P	178 ms on CP2	12.66	0.001	−0.36	0.78
		Lateral P	188 ms on P8	12.64	0.001	0.64	0.78
	270–276 ms	Left PO	270 ms on POz	10.56	0.008	−0.37	0.44
Arousal	112–138 ms	PO	124 ms on PO4	12.12	0.001	−0.48	0.75
		Frontal	130 ms on AFz	6.96	0.001	0.48	0.43
	164–174 ms	Midline CP	164 ms on Cz	6.11	0.001	−0.28	0.38
Valence	N/A						

Note: Details of each cluster can be found in Table S4.
Abbreviations: C, central sites; CP, centroparietal sites; H−A $\Delta\mu V$, happy minus angry faces amplitude difference; P, parietal sites; PO, parieto-occipital sites.

4.2 | The N170 Is Differentially Modulated by Perceived Arousal and Valence Depending on the Expression Contrast

N170 emotion effects were highly impacted by arousal and valence ratings, with the valence model significant for neutral–angry, neutral–fearful, and angry–fearful contrasts, and the arousal model significant for the neutral–angry, happy–angry, and angry–fearful contrasts. These results support our predictions and previous empirical work suggesting that the N170 is modulated by arousal (Almeida et al. 2016; Li et al. 2022), but we further show that it is also impacted by valence. Importantly, these effects depend on the expressions compared.

It is possible that our threat-specific (i.e., neutral–angry, neutral–fear, and angry–fearful contrasts) N170 valence results are linked to amygdala projections to the face network (Framorando et al. 2021), given the amygdala's role in determining stimulus valence (Anders et al. 2008; Kim et al. 2017) and its timing of discharge seen between 100 and 200 ms (Krolak-Salmon et al. 2004; Méndez-Bértolo et al. 2016). However, it is less likely that the amygdala plays a role in the arousal effects we report for anger-specific contrasts (i.e., neutral–angry, happy–angry, and angry–fearful). Indeed, fMRI work suggests amygdala activation occurs regardless of the emotion category (Gerber et al. 2008; Lin et al. 2020), which goes against our marked anger-specific arousal findings.

Rather, our arousal results are more likely due to structural components of the face (i.e., furrowed brow of an angry face, wide open eyes for fearful face). Specific structural action units engaged in facial expressions convey arousal, valence, and expression category-specific information (Liu et al. 2022), and could impact the N170 downstroke (P1 to N170 interval) where we found our effects and which reflect the integration of facial features (Schyns et al. 2007; Itier and Durston 2023). Thus, the N170 emotion effects may indicate the encoding of diagnostic features that convey arousal and/or valence. Note that local low-level factors like luminance and contrast likely contribute to this diagnosticity and to feature-conveyed arousal and valence.

Overall, our current stimulus-level approach suggests that past inconsistent N170 modulations by facial expressions may be related to participants' idiosyncratic perception of valence and arousal in each individual face stimulus. This conclusion is more probable than systematic subject-level individual differences (e.g., participant X “on average” rating happy as high on arousal), as exploratory correlation analysis revealed sparse results that did not align with the present stimulus-level arousal and valence results (Figure S2).

4.3 | The P2 Does Not Relate to Perceived Valence and Arousal

The P2, a component rarely investigated in emotion expression work (Schindler and Bublatzky 2020), was less positive for angry than for happy and neutral expressions (Han et al. 2021). The only rating-related effect around the P2 was in the happy–neutral contrast, where the amplitude difference during the N170-P2 interval related to valence ratings. However, this effect was short (20 ms), weak (F-values < 10) and seen only on right posterior electrodes.

Previous studies have interpreted P2 emotion effects as related to the emotional content of the stimulus (Calvo, Marrero, et al. 2013; Durston and Itier 2021; Han et al. 2021), perceived valence and arousal do not seem to modulate the P2. The P2 is sensitive to attended vs. unattended stimuli (Kanske et al. 2011) and typical vs. atypical faces (Kloth et al. 2017; Winward et al. 2022). Thus, the present P2 effects might index affect-irrelevant attentional mechanisms. As emotion-related P2 effects do not interact with task demands (Durston and Itier 2021; Itier and Durston 2023), these affect-irrelevant attentional processes must be bottom-up, such as spontaneous attentional capture related to an angry face (Calvo et al. 2014). Overall, more P2 research is needed in this field.

4.4 | The EPN Is Not Universally Sensitive to Perceived Valence and Arousal

Lastly, only two contrasts were significant on the EPN, with more negative amplitudes for happy than for fearful or neutral faces.

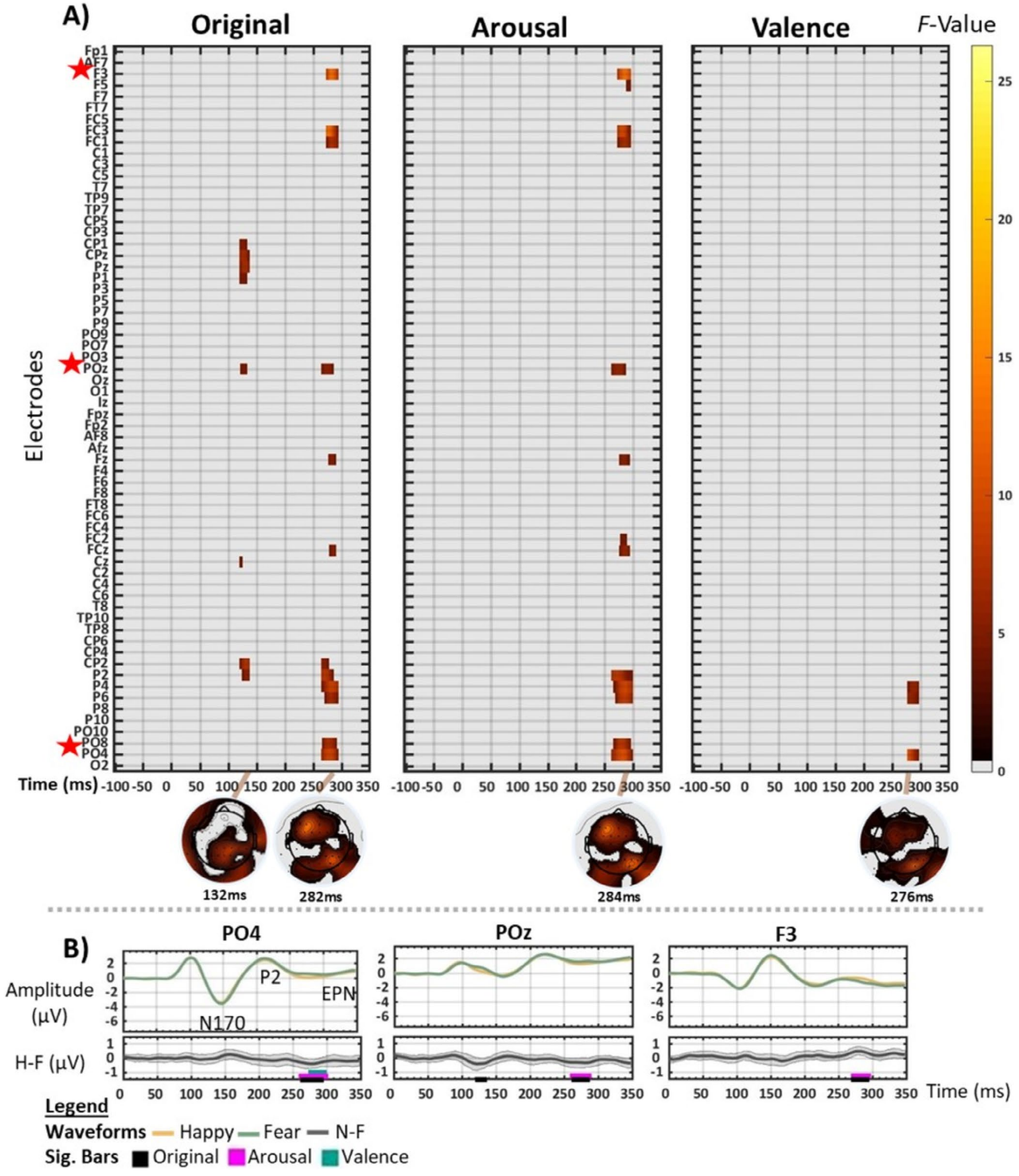

FIGURE 7 | Results from the happy and fearful expression contrasts ($p < 0.008$). (A) Time in milliseconds is along the *x*-axes and electrodes are along the *y*-axes. The color bar represents the *F*-values of significant datapoints, where gray indicates nonsignificant datapoints. *F*-values are corrected using cluster mass correction for multiple comparisons. Red stars beside electrode names indicate electrodes plotted in B. Uncorrected (i.e., not filtered using cluster mass) topographic plots underneath the raster plots represent the scalp distribution of the *F*-values at key time points where significant clusters were maximal. Topo plots use the same color scale as the rasters. (B) ERPs (top) and difference wave (bottom; happy–fearful) on key electrodes. On the difference wave plots, significance bars indicate when the contrast was significant in the original model (black), in the arousal model (magenta) and in the valence model (cyan). Confidence intervals on the difference wave plots were computed with 20% trimmed means and Bayesian bootstrapping at $\alpha = 0.008$.

TABLE 8 | Happy versus fearful *F*-contrast summary.

Model	Timing	Electrodes	Cluster max	$F(1, 79)$	p	H−F $\Delta\mu V$	D
Original	120–163 ms	Midline CP	132 ms on CP2	7.65	0.001	−0.42	0.47
	264–292 ms	Right PO	278 ms on PO4	9.38	0.001	−0.39	0.58
		Left frontal	282 ms on F3	12.87	0.001	0.48	0.79
Arousal	262–294 ms	Right PO	282 ms on PO4	10.39	0.001	−0.39	0.64
		Left frontal	284 ms on F3	13.14	0.001	0.48	0.81
Valence	276–296 ms	Right P	276 ms on PO4	10.97	0.001	−0.39	0.68

Note: Details of each cluster can be found in Table S5.
Abbreviations: CP, centroparietal sites; H−F $\Delta\mu V$, happy minus fearful faces amplitude difference; P, parietal sites; PO, parieto-occipital sites.

Our happy–neutral result replicates others (Calvo, Marrero, et al. 2013; Calvo and Beltrán 2013; Rellecke et al. 2012), including recent MU analyses (Itier and Durston 2023; but see Durston and Itier 2021). Most importantly, these marked EPN results were related to arousal *and* valence ratings, as predicted (Li et al. 2022).

Previous work suggests the EPN is involved in determining the intensity and arousal level of an expression (Calvo and Beltrán 2013; Olofsson et al. 2008). Happy faces with open mouths, as used here, are perceived as more aroused and positive than those with closed mouths and elicit more negative EPNs (Langeslag and van Strien 2018). Differences in mouth configuration across studies could thus be linked to past inconsistencies, while our stimulus-level approach likely captured the variance associated with the size of smiles from face to face. However, happy and angry faces did not differ on the EPN (Bublatzky et al. 2017; Hudson et al. 2021), a result possibly related to other factors; for example, similar self-relevance or motivational messages conveyed by the faces (Bublatzky et al. 2017; Herbert et al. 2013).

It is interesting that no other contrasts elicited significant differences on the EPN. Historically, the EPN has been *the* early attention and emotion perception component, being augmented by emotions in faces, scenes, and words (Schacht and Sommer 2009; Schupp, Junghöfer, et al. 2004; Schupp, Ohman, et al. 2004) and claimed to be sensitive to arousal, valence, and threat content (Calvo and Beltrán 2013; Olofsson et al. 2008; Schupp, Junghöfer, et al. 2004). If the EPN was *the* emotion perception component, we would expect significant differences and interactions with arousal and valence ratings in more than two contrasts. We did not find an enhanced EPN for fearful or angry compared to neutral expressions, opposing a wealth of literature (Schindler and Bublatzky 2020 for a review). The present study used high methodological and statistical controls, including MU statistics (Luck and Gaspelin 2017), a large sample ($N = 80$), and a gaze-contingent procedure (Itier and Durston 2023; Siklos-Whillans and Itier 2024). Previous work employing similar methods has found limited or no emotional effects on the EPN (Durston and Itier 2021; Hudson et al. 2021; Itier and Durston 2023). Thus, variations of the EPNs with emotional expressions reported in previous studies may have been driven by factors other than emotion per se.

4.5 | Temporal Dynamics of Facial Expression Processing

The present results suggest that affect extraction from stimuli (i.e., perceived stimuli valence and arousal) does not modulate the visual P1 and starts only after the P1, during the P1–N170 interval for threat-related expressions. In other words, structural encoding and feature integration, which are reflected in the N170 peak (Rossion and Jacques 2011) and P1–N170 interval (Itier and Durston 2023; Schyns et al. 2007; Siklos-Whillans and Itier 2024), largely overlap with the extraction of affect from threat-related expressions. This finding agrees with the view that valence and arousal are conveyed by threatening features (Liu et al. 2022). P2 amplitudes, however, likely index affect-independent attentional allocation (e.g., Kanske et al. 2011), but for angry faces only.

Interpretation of positive emotional affect, in contrast, occurs later, around the EPN, and likely involves more elaborate cognitive processes. Although previous work has suggested that the EPN is sensitive to threat-related stimuli (Schupp, Junghöfer, et al. 2004; Schupp, Ohman, et al. 2004), the present data suggest the EPN is sensitive to positive affective information. Together with the N170 findings, these data fit better with the long-standing view that threat-related expressions are prioritized temporally by the visual system (Bertini and Làdavas 2021; Calvo and Nummenmaa 2016; Ohman 2009), which contrasts with the temporal dynamics of facial expression decoding as reflected by the majority of previous ERP findings (Schindler and Bublatzky 2020 for a review). Furthermore, individual differences in average perceived valence correlate with amplitudes elicited by happy faces on the EPN (Figure S2), suggesting a potential influence of individual differences at later processing stages.

To re-iterate, the present study proposes a framework of neural processing suggesting that affect extraction from stimuli (i.e., perceived stimuli valence and arousal) occurs on the N170 for threat-related expressions and later (EPN) for positive emotions.

4.6 | Analysis Considerations

ERP data were analyzed in a categorical fashion, directly comparing predefined facial expression categories, as done by most

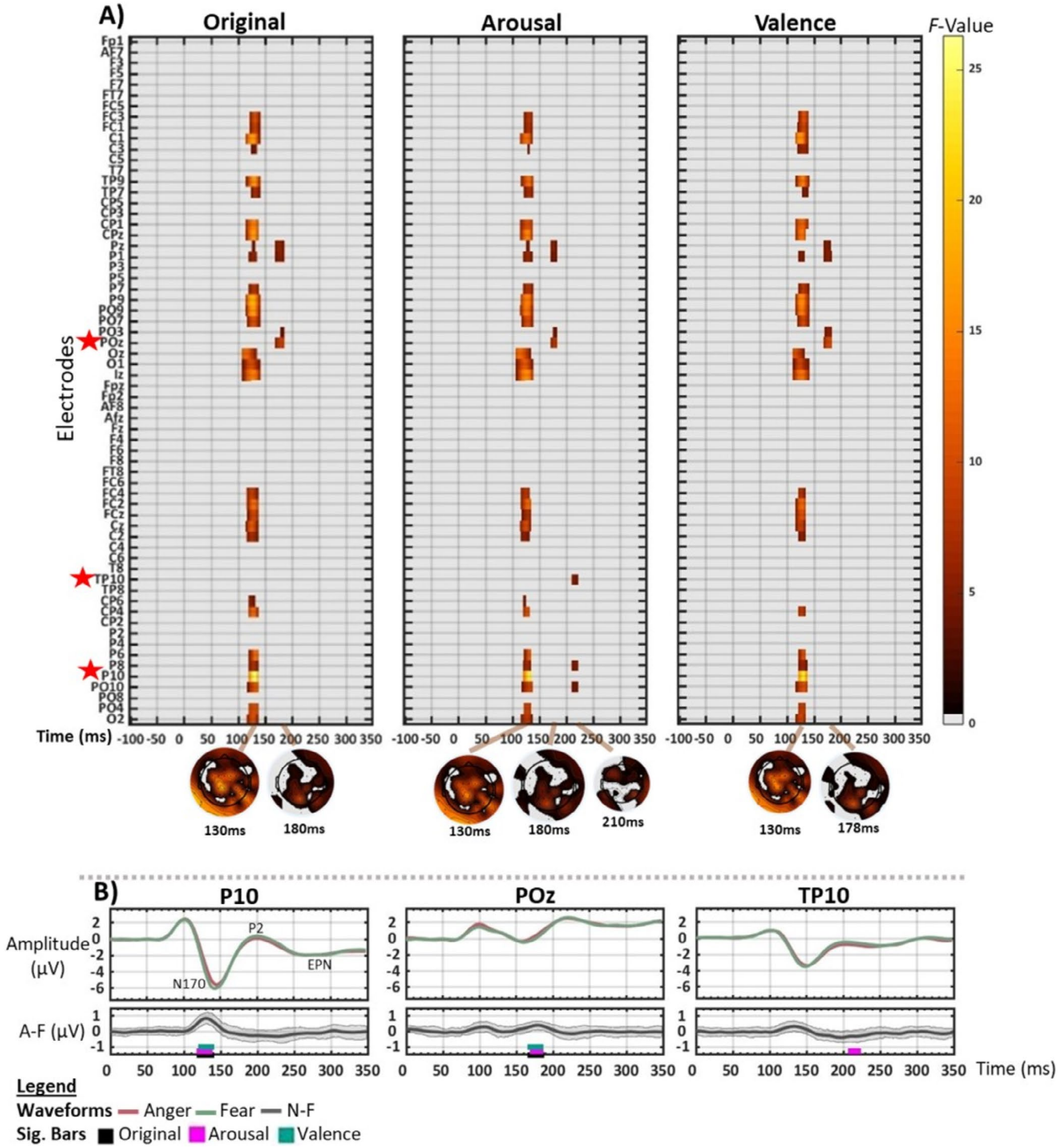

FIGURE 8 | Results from the angry and fearful expression contrasts ($p < 0.008$). (A) Time in milliseconds is along the *x*-axes and electrodes are along the *y*-axes. The color bar represents the *F*-values of significant datapoints, where gray indicates non-significant datapoints. *F*-values are corrected using cluster mass correction for multiple comparisons. Red stars beside electrode names indicate electrodes plotted in B. Uncorrected (i.e., not filtered using cluster mass) topographic plots underneath the raster plots represent the scalp distribution of the *F*-values at key time points where significant clusters were maximal. Topo plots use the same color scale as the rasters. (B) ERPs (top) and difference wave (bottom; angry–fearful) on key electrodes. On the difference wave plots, significance bars indicate when the contrast was significant in the original model (black), in the arousal model (magenta) and in the valence model (cyan). Confidence intervals on the difference wave plots were computed with 20% trimmed means and Bayesian bootstrapping at $\alpha = 0.008$.

ERP studies in the field. Valence and arousal ratings were used as regressors at the stimulus level to see how the perceived affective content of individual faces may relate to categorical expression differences on ERPs, as we believe stimulus variations within the expression category (Tracy and Robins 2007) and between databases (Adolph and Alpers 2010) must be accounted for. However, we caution readers not to interpret these results in terms of individual differences in affective ratings, as there

TABLE 9 | Angry versus fearful *F*-contrast summary.

Model	Timing	Electrodes	Cluster max	$F(1, 79)$	p	A–F $\Delta\mu V$	D
Original	108–140 ms	F, C, PO	130 ms on P10	26.30	0.001	0.83	26.30
	168–184 ms	Left & midline PO	180 ms on Poz	9.40	0.001	0.39	9.40
Arousal	106–138 ms	F, C, PO	130 ms on P10	23.92	0.001	0.83	1.48
	172–182 ms	Left & midline PO	180 ms on POz	8.09	0.005	0.39	0.50
	210–220 ms	Right lateral TP	210 ms on TP10	6.08	0.007	−0.31	0.38
Valence	112–140 ms	F, C, PO	130 ms on P10	23.84	0.001	0.83	1.47
	168–182 ms	Left & midline PO	178 ms on Poz	9.98	0.001	0.39	0.62

Note: Details of each cluster can be found in Table S6.
Abbreviations: A–F $\Delta\mu V$, angry minus fearful faces amplitude difference; C, central sites; F, frontal sites; PO, parieto-occipital sites; TP, temporoparietal sites.

are currently no methods for isolating variance from the trial and participant levels in a mass univariate analysis. Our best attempt at this was our correlational analyses (Figure S2), finding no overlap with the stimulus-based models reported. Overall, multilevel modeling (MLM) approaches would be well suited for datasets of this type. Some have begun using MLM in ERP data (e.g., Volpert-Esmond et al. 2021); however, these analyses zoom in on time points and electrodes, potentially inflating Type I and II errors (Luck and Gaspelin 2017). We encourage the field to consider full-scalp MLM analyses in the future.

Although the present study advances our understanding of the neural encoding of facial expressions, some limitations must be acknowledged. First, our arousal and valence ratings were stimulus-level, not trial-level. Participants rated faces on arousal and valence after their ERPs were recorded in a separate task, not after each trial during the ERP task. Future research should get participants to rate expressions after each trial as the ERPs are recorded (e.g., preregistered replication: https://osf.io/z7fv4). Furthermore, static images were included, yet the field is moving toward dynamic visual stimuli (e.g., Recio et al. 2014). Although we used multiple stimuli databases and highly controlled stimuli, our work should be replicated using dynamic stimuli to determine the generalizability of the present effects. Lastly, our sample was culturally diverse, with less than a third of our participants being of the same race as our stimuli (i.e., Caucasian). There is a wealth of literature surrounding how race impacts face perception (e.g., Rhodes et al. 2009) and emotions (e.g., Hu et al. 2017). However, in the present affect perception approach context, our large multicultural sample is an asset, ensuring the generalizability of results.

Despite these constraints, the present study is one of the first to conduct robust data-driven full-scalp analyses on a large sample ($N = 80$) of participants. This is also the first study to consider the relationship between arousal and valence ratings and ERP waveforms at the stimulus level in the same participants. The present results suggest that no singular ERP component is sensitive to valence or arousal. Rather, modulations of the various ERPs by affective content are contrast dependent. Thus, future research must consider the relationship between valence/arousal ratings and ERP amplitudes at the stimulus, trial, and individual levels to fully understand when and how the brain extracts affective information from faces.

Author Contributions

Amie J. Durston: conceptualization, formal analysis, methodology, visualization, writing – original draft, writing – review and editing. **Roxane J. Itier:** conceptualization, funding acquisition, methodology, resources, supervision, writing – review and editing.

Acknowledgments

Thank you to Alicia Mathew, Calla Mueller, Lily Laevens, Layla Hussain, and Suevin Un for help collecting these datum. A previous version of this manuscript was submitted for the fulfillment of the Master of Arts received by AJD. AJD was supported by NSERC and OGS scholarships.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

Final analyses output and detailed analyses steps are available on OSF (https://osf.io/zyh3e/). The individual participant data collected for this study are available from the corresponding author upon reasonable request.

Endnotes

[1] Root mean square (RMS) contrast (SD) = 0.139 (0.001) and pixel intensity (SD) = 0.933 (0.000).

[2] Additional questionnaires included STICSA (Ree et al. 2008); STAI (Spielberger 1983); SPIN (Connor et al. 2000); TRIPM (Patrick et al. 2009); LSRP (Levenson et al. 1995), not reported here.

[3] Number of trials rejected due to experimental error on average (with standard deviation): Neutral = 21.91 (17.92), Happy = 19.48 (16.42), Angry = 22.98 (19.87), Fearful = 24.76 (21.60).

[4] Number of final clean trials on average (with standard deviation): Neutral = 86.18 (20.09), Happy = 89.15 (19.76), Angry = 86.00 (22.30), Fearful = 84.00 (24.00).

[5] No model including both arousal and valence was used, as all that was highlighted was where the two models shared significant data points, which we already have by simply comparing the significant points for the valence model with the significant points for the arousal model, see result graphs.

[6] Average computed from Calvo et al. (2016), Calvo, Gutiérrez-García, et al. (2013), Calvo, Marrero, et al. (2013), Herbert et al. (2013).

References

Adolph, D., and G. W. Alpers. 2010. "Valence and Arousal: A Comparison of Two Sets of Emotional Facial Expressions." *American Journal of Psychology* 123, no. 2: 209–219. https://doi.org/10.5406/amerjpsyc.123.2.0209.

Almeida, P. R., F. Ferreira-Santos, P. L. Chaves, T. O. Paiva, F. Barbosa, and J. Marques-Teixeira. 2016. "Perceived Arousal of Facial Expressions of Emotion Modulates the N170, Regardless of Emotional Category: Time Domain and Time–Frequency Dynamics." *International Journal of Psychophysiology* 99: 48–56. https://doi.org/10.1016/j.ijpsycho.2015.11.017.

Anders, S., F. Eippert, N. Weiskopf, and R. Veit. 2008. "The Human Amygdala Is Sensitive to the Valence of Pictures and Sounds Irrespective of Arousal: An fMRI Study." *Social Cognitive and Affective Neuroscience* 3, no. 3: 233–243. https://doi.org/10.1093/scan/nsn017.

Batty, M., and M. J. Taylor. 2003. "Early Processing of the Six Basic Facial Emotional Expressions." *Cognitive Brain Research* 17, no. 3: 613–620. https://doi.org/10.1016/S0926-6410(03)00174-5.

Bertini, C., and E. Làdavas. 2021. "Fear-Related Signals Are Prioritised in Visual, Somatosensory and Spatial Systems." *Neuropsychologia* 150: 107698. https://doi.org/10.1016/j.neuropsychologia.2020.107698.

Bublatzky, F., A. Pittig, H. T. Schupp, and G. W. Alpers. 2017. "Face-To-Face: Perceived Personal Relevance Amplifies Face Processing." *Social Cognitive and Affective Neuroscience* 12, no. 5: 811–822. https://doi.org/10.1093/scan/nsx001.

Calvo, M. G., P. Avero, A. Fernández-Martín, and G. Recio. 2016. "Recognition Thresholds for Static and Dynamic Emotional Faces." *Emotion* 16, no. 8: 1186–1200. https://doi.org/10.1037/emo0000192.

Calvo, M. G., and D. Beltrán. 2013. "Recognition Advantage of Happy Faces: Tracing the Neurocognitive Processes." *Neuropsychologia* 51, no. 11: 2051–2061. https://doi.org/10.1016/j.neuropsychologia.2013.07.010.

Calvo, M. G., A. Gutiérrez-García, P. Avero, and D. Lundqvist. 2013a. "Attentional Mechanisms in Judging Genuine and Fake Smiles: Eye-Movement Patterns." *Emotion* 13, no. 4: 792–802. https://doi.org/10.1037/a0032317.

Calvo, M. G., A. Gutiérrez-García, A. Fernández-Martín, and L. Nummenmaa. 2014. "Recognition of Facial Expressions of Emotion Is Related to Their Frequency in Everyday Life." *Journal of Nonverbal Behavior* 38, no. 4: 549–567. https://doi.org/10.1007/s10919-014-0191-3.

Calvo, M. G., H. Marrero, and D. Beltrán. 2013b. "When Does the Brain Distinguish Between Genuine and Ambiguous Smiles? An ERP Study." *Brain and Cognition* 81, no. 2: 237–246. https://doi.org/10.1016/j.bandc.2012.10.009.

Calvo, M. G., and L. Nummenmaa. 2016. "Perceptual and Affective Mechanisms in Facial Expression Recognition: An Integrative Review." *Cognition & Emotion* 30, no. 6: 1081–1106. https://doi.org/10.1080/02699931.2015.1049124.

Connor, K. M., J. R. T. Davidson, L. E. Churchill, A. Sherwood, R. H. Weisler, and E. Foa. 2000. "Psychometric Properties of the Social Phobia Inventory (SPIN): New Self-Rating Scale." *British Journal of Psychiatry* 176, no. 4: 379–386. https://doi.org/10.1192/bjp.176.4.379.

Delorme, A., and S. Makeig. 2004. "EEGLAB: An Open Source Toolbox for Analysis of Single-Trial EEG Dynamics Including Independent Component Analysis." *Journal of Neuroscience Methods* 134, no. 1: 9–21. https://doi.org/10.1016/j.jneumeth.2003.10.009.

Durston, A. J., and R. J. Itier. 2021. "The Early Processing of Fearful and Happy Facial Expressions Is Independent of Task Demands—Support From Mass Univariate Analyses." *Brain Research* 1765: 147505. https://doi.org/10.1016/j.brainres.2021.147505.

Ebner, N. C., M. Riediger, and U. Lindenberger. 2010. "FACES—A Database of Facial Expressions in Young, Middle-Aged, and Older Women and Men: Development and Validation." *Behavior Research Methods* 42, no. 1: 351–362. https://doi.org/10.3758/BRM.42.1.351.

Framorando, D., E. Moses, L. Legrand, M. Seeck, and A. J. Pegna. 2021. "Rapid Processing of Fearful Faces Relies on the Right Amygdala: Evidence From Individuals Undergoing Unilateral Temporal Lobectomy." *Scientific Reports* 11: 426. https://doi.org/10.1038/s41598-020-80054-1.

Gerber, A. J., J. Posner, D. Gorman, et al. 2008. "An Affective Circumplex Model of Neural Systems Subserving Valence, Arousal, and Cognitive Overlay During the Appraisal of Emotional Faces." *Neuropsychologia* 46, no. 8: 2129–2139. https://doi.org/10.1016/j.neuropsychologia.2008.02.032.

Han, S., J. Hu, W. Li, et al. 2021. "From Structure to Concepts: The Two Stages of Facial Expression Recognition." *Neuropsychologia* 150: 107700. https://doi.org/10.1016/j.neuropsychologia.2020.107700.

Herbert, C., A. Sfaerlea, and T. Blumenthal. 2013. "Your Emotion or Mine: Labeling Feelings Alters Emotional Face Perception—An ERP Study on Automatic and Intentional Affect Labeling." *Frontiers in Human Neuroscience* 7: 378. https://doi.org/10.3389/fnhum.2013.00378.

Hinojosa, J. A., F. Mercado, and L. Carretié. 2015. "N170 Sensitivity to Facial Expression: A Meta-Analysis." *Neuroscience & Biobehavioral Reviews* 55: 498–509. https://doi.org/10.1016/j.neubiorev.2015.06.002.

Hu, C. S., Q. Wang, T. Han, E. Weare, and G. Fu. 2017. "Differential Emotion Attribution to Neutral Faces of Own and Other Races." *Cognition and Emotion* 31, no. 2: 360–368. https://doi.org/10.1080/02699931.2015.1092419.

Hudson, A., A. J. Durston, S. D. McCrackin, and R. J. Itier. 2021. "Emotion, Gender and Gaze Discrimination Tasks Do Not Differentially Impact the Neural Processing of Angry or Happy Facial Expressions—A Mass Univariate ERP Analysis." *Brain Topography* 34, no. 6: 813–833. https://doi.org/10.1007/s10548-021-00873-x.

Itier, R. J., and A. J. Durston. 2023. "Mass-Univariate Analysis of Scalp ERPs Reveals Large Effects of Gaze Fixation Location During Face Processing That Only Weakly Interact With Face Emotional Expression." *Scientific Reports* 13, no. 1: 17022. https://doi.org/10.1038/s41598-023-44355-5.

Itier, R. J., and K. N. Neath-Tavares. 2017. "Effects of Task Demands on the Early Neural Processing of Fearful and Happy Facial Expressions." *Brain Research* 1663: 38–50. https://doi.org/10.1016/j.brainres.2017.03.013.

Kanske, P., J. Plitschka, and S. A. Kotz. 2011. "Attentional Orienting Towards Emotion: P2 and N400 ERP Effects." *Neuropsychologia* 49, no. 11: 3121–3129. https://doi.org/10.1016/j.neuropsychologia.2011.07.022.

Keltner, D., D. Sauter, J. Tracy, and A. Cowen. 2019. "Emotional Expression: Advances in Basic Emotion Theory." *Journal of Nonverbal Behavior* 43, no. 2: 133–160. https://doi.org/10.1007/s10919-019-00293-3.

Kim, M. J., A. M. Mattek, R. H. Bennett, K. M. Solomon, J. Shin, and P. J. Whalen. 2017. "Human Amygdala Tracks a Feature-Based Valence Signal Embedded Within the Facial Expression of Surprise." *Journal of Neuroscience* 37, no. 39: 9510–9518. https://doi.org/10.1523/JNEUROSCI.1375-17.2017.

Kline, R. B. 1998. *Principles and Practice of Structural Equation Modeling*. Guilford Press.

Kloth, N., G. Rhodes, and S. R. Schweinberger. 2017. "Watching the Brain Recalibrate: Neural Correlates of Renormalization During Face Adaptation." *NeuroImage* 155: 1–9. https://doi.org/10.1016/j.neuroimage.2017.04.049.

Krolak-Salmon, P., M.-A. Hénaff, A. Vighetto, O. Bertrand, and F. Mauguière. 2004. "Early Amygdala Reaction to Fear Spreading in Occipital, Temporal, and Frontal Cortex: A Depth Electrode ERP Study in Human." *Neuron* 42, no. 4: 665–676. https://doi.org/10.1016/S0896-6273(04)00264-8.

Langeslag, S. J. E., and J. W. van Strien. 2018. "Early Visual Processing of Snakes and Angry Faces: An ERP Study." *Brain Research* 1678: 297–303. https://doi.org/10.1016/j.brainres.2017.10.031.

Langner, O., R. Dotsch, G. Bijlstra, D. H. J. Wigboldus, S. T. Hawk, and A. van Knippenberg. 2010. "Presentation and Validation of the Radboud Faces Database." *Cognition and Emotion* 24, no. 8: 1377–1388. https://doi.org/10.1080/02699930903485076.

Lee, M. D., and E. J. Wagenmakers. 2014. *Bayesian Cognitive Modeling: A Practical Course*. Cambridge University Press.

Levenson, M. R., K. A. Kiehl, and C. M. Fitzpatrick. 1995. "Assessing Psychopathic Attributes in a Noninstitutionalized Population." *Journal of Personality and Social Psychology* 68, no. 1: 151–158. https://doi.org/10.1037//0022-3514.68.1.151.

Li, Y., M. Zhang, S. Liu, and W. Luo. 2022. "EEG Decoding of Multidimensional Information From Emotional Faces." *NeuroImage* 258: 119374. https://doi.org/10.1016/j.neuroimage.2022.119374.

Lin, H., M. Müller-Bardorff, B. Gathmann, et al. 2020. "Stimulus Arousal Drives Amygdalar Responses to Emotional Expressions Across Sensory Modalities." *Scientific Reports* 10, no. 1: 1898. https://doi.org/10.1038/s41598-020-58839-1.

Liu, M., Y. Duan, R. A. A. Ince, et al. 2022. "Facial Expressions Elicit Multiplexed Perceptions of Emotion Categories and Dimensions." *Current Biology* 32, no. 1: 200–209.e6. https://doi.org/10.1016/j.cub.2021.10.035.

Lopez-Calderon, J., and S. J. Luck. 2014. "ERPLAB: An Open-Source Toolbox for the Analysis of Event-Related Potentials." *Frontiers in Human Neuroscience* 8: 213. https://doi.org/10.3389/fnhum.2014.00213.

Luck, S. J., and N. Gaspelin. 2017. "How to Get Statistically Significant Effects in any ERP Experiment (And Why You Shouldn't)." *Psychophysiology* 54, no. 1: 146–157. https://doi.org/10.1111/psyp.12639.

Ma, D. S., J. Correll, and B. Wittenbrink. 2015. "The Chicago Face Database: A Free Stimulus Set of Faces and Norming Data." *Behavior Research Methods* 47, no. 4: 1122–1135. https://doi.org/10.3758/s13428-014-0532-5.

Maris, E., and R. Oostenveld. 2007. "Nonparametric Statistical Testing of EEG- and MEG-Data." *Journal of Neuroscience Methods* 164, no. 1: 177–190. https://doi.org/10.1016/j.jneumeth.2007.03.024.

Méndez-Bértolo, C., S. Moratti, R. Toledano, et al. 2016. "A Fast Pathway for Fear in Human Amygdala." *Nature Neuroscience* 19, no. 8: 1041–1049. https://doi.org/10.1038/nn.4324.

Müller-Bardorff, M., M. Bruchmann, M. Mothes-Lasch, et al. 2018. "Early Brain Responses to Affective Faces: A Simultaneous EEG-fMRI Study." *NeuroImage* 178: 660–667. https://doi.org/10.1016/j.neuroimage.2018.05.081.

Ohman, A. 2009. "Of Snakes and Faces: An Evolutionary Perspective on the Psychology of Fear." *Scandinavian Journal of Psychology* 50, no. 6: 543–552. https://doi.org/10.1111/j.1467-9450.2009.00784.x.

Olofsson, J. K., S. Nordin, H. Sequeira, and J. Polich. 2008. "Affective Picture Processing: An Integrative Review of ERP Findings." *Biological Psychology* 77, no. 3: 247–265. https://doi.org/10.1016/j.biopsycho.2007.11.006.

Palermo, R., and G. Rhodes. 2007. "Are You Always on My Mind? A Review of How Face Perception and Attention Interact." *Neuropsychologia* 45, no. 1: 75–92. https://doi.org/10.1016/j.neuropsychologia.2006.04.025.

Patrick, C. J., D. C. Fowles, and R. F. Krueger. 2009. "Triarchic Conceptualization of Psychopathy: Developmental Origins of Disinhibition, Boldness, and Meanness." *Development and Psychopathology* Cambridge Core 21, no. 3: 913–938. https://doi.org/10.1017/S0954579409000492.

Pernet, C. R., N. Chauveau, C. Gaspar, and G. A. Rousselet. 2011. "LIMO EEG: A Toolbox for Hierarchical LInear MOdeling of ElectroEncephaloGraphic Data." *Computational Intelligence and Neuroscience* 2011: 831409. https://doi.org/10.1155/2011/831409.

Pernet, C. R., M. Latinus, T. E. Nichols, and G. A. Rousselet. 2015. "Cluster-Based Computational Methods for Mass Univariate Analyses of Event-Related Brain Potentials/Fields: A Simulation Study." *Journal of Neuroscience Methods* 250: 85–93. https://doi.org/10.1016/j.jneumeth.2014.08.003.

Pernet, C. R., R. Martinez-Cancino, D. Truong, S. Makeig, and A. Delorme. 2021. "From BIDS-Formatted EEG Data to Sensor-Space Group Results: A Fully Reproducible Workflow With EEGLAB and LIMO EEG." *Frontiers in Neuroscience* 14: 610388. https://doi.org/10.3389/fnins.2020.610388.

Qiu, Z., S. I. Becker, H. Xia, Z. Hamblin-Frohman, and A. J. Pegna. 2023. "Fixation-Related Electrical Potentials During a Free Visual Search Task Reveal the Timing of Visual Awareness." *iScience* 26, no. 7: 107148. https://doi.org/10.1016/j.isci.2023.107148.

Recio, G., A. Schacht, and W. Sommer. 2014. "Recognizing Dynamic Facial Expressions of Emotion: Specificity and Intensity Effects in Event-Related Brain Potentials." *Biological Psychology* 96: 111–125. https://doi.org/10.1016/j.biopsycho.2013.12.003.

Ree, M., D. French, C. MacLeod, and V. Locke. 2008. "Distinguishing Cognitive and Somatic Dimensions of State and Trait Anxiety: Development and Validation of the State-Trait Inventory for Cognitive and Somatic Anxiety (STICSA)." *Behavioural and Cognitive Psychotherapy* 36: 313–332. https://doi.org/10.1017/S1352465808004232.

Rellecke, J., W. Sommer, and A. Schacht. 2012. "Does Processing of Emotional Facial Expressions Depend on Intention? Time-Resolved Evidence From Event-Related Brain Potentials." *Biological Psychology* 90, no. 1: 23–32. https://doi.org/10.1016/j.biopsycho.2012.02.002.

Rhodes, G., V. Locke, L. Ewing, and E. Evangelista. 2009. "Race Coding and the Other-Race Effect in Face Recognition." *Perception* 38, no. 2: 232–241. https://doi.org/10.1068/p6110.

Rossion, B., and C. Jacques. 2011. "The N170: Understanding the Time Course of Face Perception in the Human Brain." In *The Oxford Handbook of ERP Components*, 115–142. https://doi.org/10.1093/oxfordhb/9780195374148.013.0064.

Schacht, A., and W. Sommer. 2009. "Emotions in Word and Face Processing: Early and Late Cortical Responses." *Brain and Cognition* 69, no. 3: 538–550. https://doi.org/10.1016/j.bandc.2008.11.005.

Schindler, S., and F. Bublatzky. 2020. "Attention and Emotion: An Integrative Review of Emotional Face Processing as a Function of Attention." *Cortex* 130: 362–386. https://doi.org/10.1016/j.cortex.2020.06.010.

Schindler, S., S. Meyer, M. Bruchmann, N. A. Busch, and T. Straube. 2023. "Fearful Faces Straight Ahead or in the Periphery: Early Neuronal Responses Independently of Trait Anxiety." *Emotion* 23, no. 6: 1687–1701. https://doi.org/10.1037/emo0001184.

Schupp, H. T., M. Junghöfer, A. I. Weike, and A. O. Hamm. 2004. "The Selective Processing of Briefly Presented Affective Pictures: An ERP Analysis." *Psychophysiology* 41, no. 3: 441–449. https://doi.org/10.1111/j.1469-8986.2004.00174.x.

Schupp, H. T., A. Ohman, M. Junghöfer, A. I. Weike, J. Stockburger, and A. O. Hamm. 2004. "The Facilitated Processing of Threatening Faces: An ERP Analysis." *Emotion* 4, no. 2: 189–200. https://doi.org/10.1037/1528-3542.4.2.189.

Schyns, P. G., L. S. Petro, and M. L. Smith. 2007. "Dynamics of Visual Information Integration in the Brain for Categorizing Facial Expressions." *Current Biology* 17, no. 18: 1580–1585. https://doi.org/10.1016/j.cub.2007.08.048.

Siklos-Whillans, J., and R. J. Itier. 2024. "Effects of Inversion and Fixation Location on the Processing of Face and House Stimuli—A Mass Univariate Analysis." *Brain Topography* 37: 972–992. https://doi.org/10.1007/s10548-024-01068-w.

Smith, E., A. Weinberg, T. Moran, and G. Hajcak. 2013. "Electrocortical Responses to NIMSTIM Facial Expressions of Emotion." *International Journal of Psychophysiology* 88, no. 1: 17–25. https://doi.org/10.1016/j.ijpsycho.2012.12.004.

Spielberger, C. D. 1983. *Manual for the State-Trait Anxiety Inventory*. Mind Garden.

Tottenham, N., J. W. Tanaka, A. C. Leon, et al. 2009. "The NimStim Set of Facial Expressions: Judgments From Untrained Research Participants." *Psychiatry Research* 168, no. 3: 242–249. https://doi.org/10.1016/j.psychres.2008.05.006.

Tracy, J. L., and R. W. Robins. 2007. "The Prototypical Pride Expression: Development of a Nonverbal Behavior Coding System." *Emotion* 7, no. 4: 789–801. https://doi.org/10.1037/1528-3542.7.4.789.

Turano, M. T., J. Lao, A.-R. Richoz, et al. 2017. "Fear Boosts the Early Neural Coding of Faces." *Social Cognitive and Affective Neuroscience* 12, no. 12: 1959–1971. https://doi.org/10.1093/scan/nsx110.

Volpert-Esmond, H. I., E. Page-Gould, and B. D. Bartholow. 2021. "Using Multilevel Models for the Analysis of Event-Related Potentials." *International Journal of Psychophysiology* 162: 145–156. https://doi.org/10.1016/j.ijpsycho.2021.02.006.

Vuilleumier, P., and G. Pourtois. 2007. "Distributed and Interactive Brain Mechanisms During Emotion Face Perception: Evidence From Functional Neuroimaging." *Neuropsychologia* 45, no. 1: 174–194. https://doi.org/10.1016/j.neuropsychologia.2006.06.003.

Willenbockel, V., J. Sadr, D. Fiset, G. O. Horne, F. Gosselin, and J. W. Tanaka. 2010. "Controlling Low-Level Image Properties: The SHINE Toolbox." *Behavior Research Methods* 42, no. 3: 671–684. https://doi.org/10.3758/BRM.42.3.671.

Williams, L. M., D. Palmer, B. J. Liddell, L. Song, and E. Gordon. 2006. "The 'When' and 'Where' of Perceiving Signals of Threat Versus Non-threat." *NeuroImage* 31, no. 1: 458–467. https://doi.org/10.1016/j.neuroimage.2005.12.009.

Winward, S. B., J. Siklos-Whillans, and R. J. Itier. 2022. "Impact of Face Outline, Parafoveal Feature Number and Feature Type on Early Face Perception in a Gaze-Contingent Paradigm: A Mass-Univariate Re-Analysis of ERP Data." *Neuroimage: Reports* 2, no. 4: 100148. https://doi.org/10.1016/j.ynirp.2022.100148.

Supporting Information

Additional supporting information can be found online in the Supporting Information section.