NeuroImage 179 (2018) 235–251

Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/neuroimage

Tuning functions for automatic detection of brief changes of facial expression in the human brain

Arnaud Leleu [a,*], Milena Dzhelyova [b], Bruno Rossion [b,c,d], Renaud Brochard [a], Karine Durand [a], Benoist Schaal [a], Jean-Yves Baudouin [a,e]

[a] Centre des Sciences Du Goût et de L'Alimentation, AgroSup Dijon, CNRS, INRA, Université Bourgogne Franche-Comté, F-21000, Dijon, France
[b] Psychological Sciences Research Institute, Institute of Neuroscience, University of Louvain, 1348, Louvain-la-Neuve, Belgium
[c] Université de Lorraine, CNRS, CRAN, F-54000, Nancy, France
[d] Université de Lorraine, CHRU-Nancy, Service de Neurologie, F-54000, Nancy, France
[e] Laboratoire Développement, Individu, Processus, Handicap, Éducation (DIPHE), Département Psychologie du Développement, de l'Éducation et des Vulnérabilités (PsyDÉV), Institut de psychologie, Université de Lyon (Lumière Lyon 2), 69676 Bron cedex, France

ARTICLE INFO

Keywords:
EEG
FPVS
Sweep-VEP
Facial expression
Intensity
Tuning function

ABSTRACT

Efficient decoding of even brief and slight intensity facial expression changes is important for social interactions. However, robust evidence for the human brain ability to automatically detect brief and subtle changes of facial expression remains limited. Here we built on a recently developed paradigm in human electrophysiology with full-blown expressions (Dzhelyova et al., 2017), to isolate and quantify a neural marker for the detection of brief and subtle changes of facial expression. Scalp electroencephalogram (EEG) was recorded from 18 participants during stimulation of a neutral face changing randomly in size at a rapid rate of 6 Hz. Brief changes of expression appeared every five stimulation cycle (i.e., at 1.2 Hz) and expression intensity increased parametrically every 20 s in 20% steps during sweep sequences of 100 s. A significant 1.2 Hz response emerged in the EEG spectrum already at 40% of facial expression-change intensity for most of the 5 emotions tested (anger, disgust, fear, happiness, or sadness in different sequences), and increased with intensity steps, predominantly over right occipito-temporal regions. Given the high signal-to-noise ratio of the approach, thresholds for automatic detection of brief changes of facial expression could be determined for every single individual brain. A time-domain analysis revealed three components, the two first increasing linearly with increasing intensity as early as 100 ms after a change of expression, suggesting gradual low-level image-change detection prior to visual coding of facial movements. In contrast, the third component showed abrupt sensitivity to increasing expression intensity beyond 300 ms post expression-change, suggesting categorical emotion perception. Overall, this characterization of the detection of subtle changes of facial expression and its temporal dynamics open promising tracks for precise assessment of social perception ability during development and in clinical populations.

Introduction

During social interactions, humans communicate a wealth of information through non-verbal behavior, among which facial expression of emotions constitutes a critical cue to infer the affective states of conspecifics and adjust behavior. Charles Darwin suggested in his seminal work on the expression of emotions that the ability to recognize emotional expressions has been selected during evolution to increase survival in social groups (Darwin, 1872). Following this rationale, Paul Ekman considered six basic emotions eliciting specific patterns of facial actions universally categorized by humans: anger, disgust, happiness, fear, sadness and surprise (e.g., Ekman, 1992). Despite some variations (Jack et al., 2012; Miellet et al., 2013), it is generally acknowledged that these facial expressions of emotions are quite well-recognized across cultures (Elfenbein and Ambady, 2002 for a meta-analysis; Sauter and Eisner, 2013).

In the large body of research about facial expression perception, most studies focus on highly expressive faces. However, despite their evident contribution to our knowledge of the function, full-blown prototypical expressions are not the most frequently encountered in everyday-life

* Corresponding author. Centre des Sciences du Goût et de l'Alimentation, AgroSup Dijon, UMR 6265 CNRS, UMR 1324 INRA, Université Bourgogne Franche-Comté, 9E Boulevard Jeanne d'Arc, 21000, Dijon, France.
E-mail address: arnaud.leleu@u-bourgogne.fr (A. Leleu).

https://doi.org/10.1016/j.neuroimage.2018.06.048
Received 5 October 2017; Received in revised form 3 May 2018; Accepted 15 June 2018
Available online 18 June 2018

(Horstmann, 2002; Motley and Camden, 1988), and their exaggerated nature has long been criticized by some authors (e.g., Carroll and Russell, 1997). Delineating our sensitivity to subtle facial cues is therefore critical for a comprehensive understanding of facial expression perception in both healthy and pathological participants (Calder et al., 2000a; Calvo et al., 2016; Etcoff and Magee, 1992; Gao and Maurer, 2010; Hess et al., 1997; Leleu et al., 2016; Marneweck et al., 2013). In particular, parametric manipulations of expression intensity by means of a linear continuum of morphs between a neutral and an expressive face allow to determine a detection threshold (i.e., the lowest intensity for accurate categorization of the expression) and the shape of a tuning function (i.e., linear vs. abrupt detection) to any facial emotional signal (e.g., Etcoff and Magee, 1992; Gao and Maurer, 2010; Hess et al., 1997; Leleu et al., 2016).

Detection thresholds are a sensitive measure of expression perception abilities. For instance, while all facial emotions are well-recognized at full intensity, only a subset are still recognized above chance level at very low intensity (e.g., happiness *vs.* other emotions, Calvo et al., 2016). During typical development, 5-year old children reach adult-like accuracy to categorize some full intensity expressions (anger, fear, sadness) but they need higher intensities than adults (Gao and Maurer, 2010). The sensitivity of detection thresholds is also particularly relevant for investigating syndromes associated with social-cognitive deficits. For example, while high-intensity happy faces are well categorized by children and adolescents with 22q11.2 deletion syndrome, they need greater intensity of happiness to reach the same accuracy than healthy participants (Leleu et al., 2016).

Regarding the shape of the tuning functions to facial expressions, behavioral studies generally found that emotions in faces are detected categorically, that is, with an abrupt perception of the emotion with increasing expression intensity (Etcoff and Magee, 1992; Leleu et al., 2016; Utama et al., 2009), consistent with other findings obtained with morphs between two emotional expressions (Calder et al., 1996; Vernet et al., 2008; Young et al., 1997). This has been shown with both identification tasks asking for verbal labeling of the emotion (Etcoff and Magee, 1992; Utama et al., 2009) and visual matching or discrimination tasks (Etcoff and Magee, 1992; Leleu et al., 2016), thus suggesting that expression categories have sharp boundaries. However, studies using an emotional intensity rating task rather found linear response profiles (Calder et al., 2000a; Hess et al., 1997; Utama et al., 2009). This dissociation between emotion categorization reaching high accuracy level through abrupt increase *vs.* linear increment of intensity ratings points to the critical influence of the explicit task performed when measuring the output of the cognitive system, a mixture of perceptual (e.g. visual categorization) and post-perceptual (e.g., motivation, decision) processes.

To overcome these issues, electroencephalographic (EEG) activity has been recorded in response to facial expressions varying in intensity. A first study tested three facial expressions (anger, disgust, fear) expressed at three intensities (50, 100, 150%) and found gradual enhancement of electrocortical activities with increasing intensity for all expressions from the N170 event-related potential (ERP) and until 600 ms post-stimulus in occipito-temporal regions (Sprengelmeyer and Jentzsch, 2006). The face-sensitive N170 was interpreted as early saliency coding favoring later visual categorization of the expressions. However, no similar effect was found for happiness in a subsequent study (Leppänen et al., 2007), suggesting that the saliency coding system is only dedicated to negative facial expressions. In contrast, another study using five nonlinear intensity levels (i.e., selected from a preliminary behavioral experiment) for happy and disgusted facial expressions showed ERP enhancement with increasing intensity already at the level of the sensory P1 component (i.e., from 90 ms after stimulus-onset) and further in the time-range of the face-sensitive N170 (i.e., 140–190 ms) for both expressions (Utama et al., 2009). Dissociation between the P1 and the N170 was found by correlating their amplitude with the data from two preliminary behavioral tasks. P1 increase was only significantly associated with abrupt categorization of facial expression, whereas the N170 intensity effect showed significant relation only with the gradual rating of expression intensity. However, strong correlations with behavioral measures were evident for both ERP responses. Moreover, the use of nonlinear increment of intensity hampered direct estimation of the abrupt *vs.* linear increase of ERP amplitude. Finally, a more recent study assessed anger detection at three linear intensity levels (i.e., 20, 60, 100%) and revealed enhanced amplitudes of both the P1 and N170 with increasing intensity (Wang et al., 2013). The authors suggested an early gain control of visual resources indexed by the P1 before the perceptual coding of facial configuration deformation at the level of the N170, but the absence of differential effects between the two components precludes strong support for a functional dissociation.

In summary, while a few ERP studies have investigated the visual processing of facial expression perception as a function of intensity, no firm conclusion could be drawn from these observations. In particular, important discrepancies remain about the advent of facial expression intensity effects throughout the processing stream, and many issues are unresolved such as the generalizability *vs.* specificity across emotion categories, the different functional mechanisms indexed by the electrocortical responses and also the shape of the tuning functions depending on intensity. These discrepancies may be partly due to the low signal-to-noise ratio (SNR) of the standard ERP approach, which requires many trials for the same condition to reach acceptable levels of SNRs for most components (Luck, 2005). Unfortunately, these long recording sessions prevent from testing various emotions expressed at various intensities in a single experiment. Moreover, subjective definitions of ERP components within more or less broad time-windows of interest reduce the generalizability of the results across studies. Besides, expression-specific activities (i.e., excluding general visual mechanisms common to both neutral and emotional expressions) are extracted by *post-hoc* subtraction of averaged 'neutral' trials from averaged 'expression' trials. This manipulation thus only indirectly captures the neural activities subtending the discrimination of an expression from neutrality (note that two studies cited above have circumvented this issue by measuring ERPs during the presentation of a facial expression that directly follows a neutral face with no blank interval in between; Utama et al., 2009; Wang et al., 2013). Finally, visual responses occurring beyond 200 ms post-stimulus are not well-captured by ERPs, due to various sources of additional noise. This is particularly the case when stimuli are – as often – presented for relatively long durations and visual exploration (i.e., leading to accumulating neural responses elicited by new retinal patterns of activity at each fixation) as well as non-perceptual processing (e.g., verbal labelling) contaminate the electrophysiological responses. This is an important point since different facial expressions elicit different patterns of visual exploration when they are presented for long durations (e.g., Bombari et al., 2013). In addition, if a behavioral response is required at each stimulus-onset, post-perceptual (e.g., decisional) processes may also contaminate the signal.

Considering these limitations, here we aimed at providing a quantified electrophysiological measure of the detection of facial expressions as a function of their intensity in the human brain with a fast periodic visual stimulation (FPVS) approach. FPVS-EEG is based on the old observation that a periodic sensory input elicits a same-frequency periodic EEG response over corresponding sensory cortices (Adrian and Matthews, 1934), the so-called *steady-state visual evoked potentials* ("SSVEP", Regan, 1989; Norcia et al., 2015 for review). This approach is particularly well suited to quantify brain responses because of its objectivity (i.e., response are measured at pre-experimentally defined frequencies) and high SNR, providing significant responses in a few minutes of recording often quantifiable in every individual participant. It has been recently adapted to investigate high-level visual processes such as generic face categorization (Rossion et al., 2015), face individualization (Liu-Shuang et al., 2014), and most importantly for the purpose of the current study the detection of brief facial expression changes (Dzhelyova et al., 2017). In the latter study, an "FPVS oddball paradigm" inspired by the well-known

oddball paradigm used in ERP studies (see sections 2.3, 4.1 and 4.3 for further discussion) was designed by inserting an individual face expressing a deviant emotion every 5th stimuli in a fast periodic train of the same neutral face repeated at 5.88 Hz. An expression-change specific EEG response was thus isolated at 5.88 Hz/5 = 1.18 Hz, directly reflecting the discrimination of the deviant facial expression from the neutral standard. Different topographical patterns of activity over occipito-parietal and occipito-temporal regions were identified for three emotional expressions (disgust, fear, happiness) and significant responses quantified in every participant. In this paradigm, the response is recorded implicitly (participants performing an orthogonal non-periodic behavioral task), it is elicited following a single glance (i.e., the fast mode of presentation displays each picture for less than 170 ms), and it is not due to low-level detection of image-based changes since it is drastically reduced by picture-plane inversion of the stimuli (Dzhelyova et al., 2017). The temporal dynamics of this discriminative response of expressive from neutral faces can also be directly established from a time-domain analysis of 5.88 Hz filtered-out EEG data.

Here we generalize these findings to a different stimulus set and extend them by testing five emotional expressions (anger, disgust, fear, happiness and sadness). Most importantly, we manipulate expression intensity in a parametric fashion using five linear levels of morphs between neutrality and expressions (20, 40, 60, 80 and 100%) and design a sweep visual evoked potentials ("sweep-VEP"; Regan, 1973; see Ales et al., 2012 for an extension to high-level visual stimuli) approach by increasing expression intensity every 20 s in 20% steps (Fig. 1). Hence, we define the tuning function to brief expression changes of increasing intensities for several emotions, as well as individual 'neural' detection thresholds (i.e., how much expression intensity is sufficient to elicit a significant expression-change response in each individual brain). By analyzing the expression-change responses in the time-domain, we also characterize the temporal dynamics of these neural tuning functions throughout the processing stream.

Material and methods

Participants

We tested eighteen participants (10 females, 2 left-handed (1 female); mean age: 25.9 ± 3.7 (SD) years, range: 21–32 years). All reported

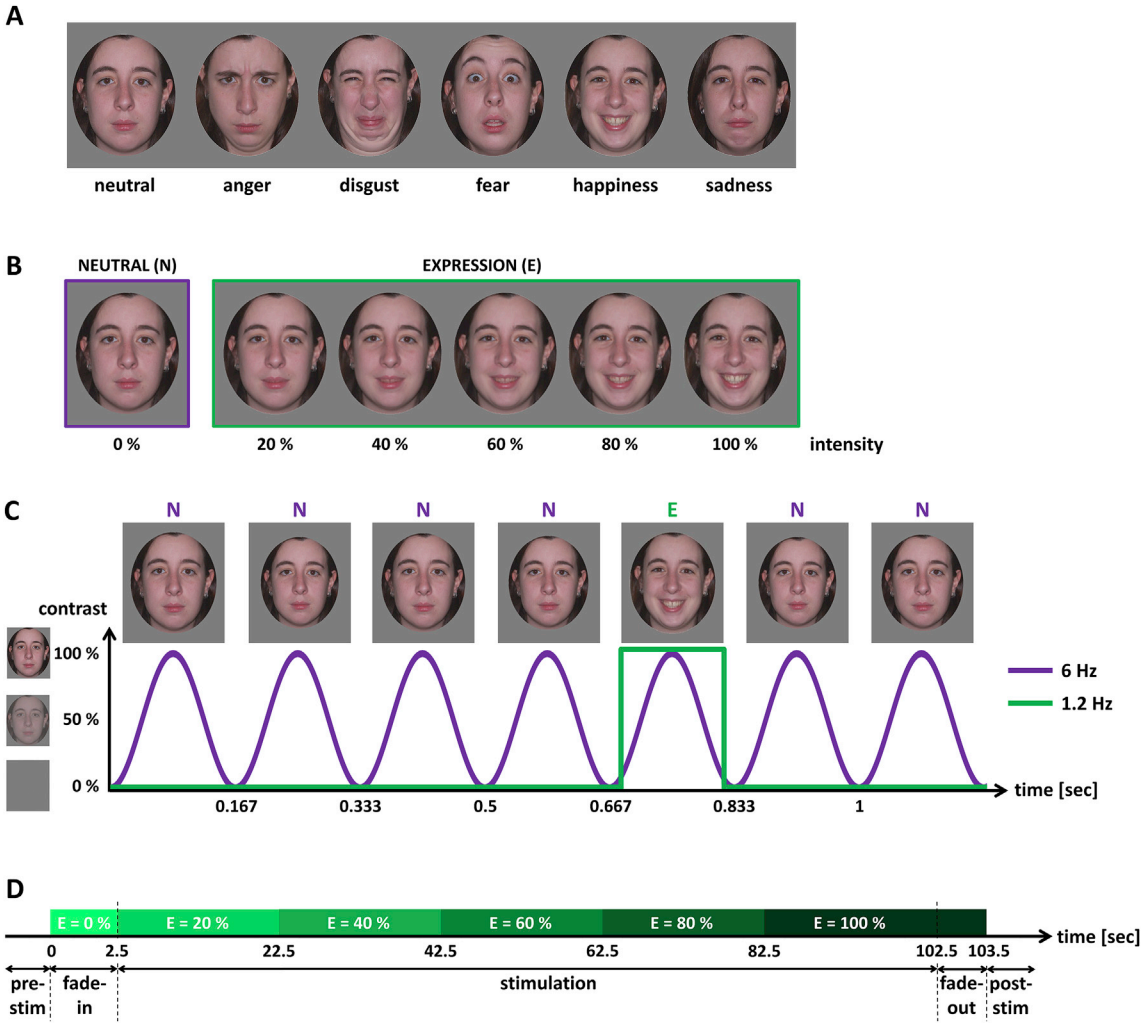


Fig. 1. The facial expression intensity sweep-VEP paradigm. A. An individual female face with a neutral expression or expressing the five emotions used in the study (anger, disgust, fear, happiness, sadness). **B.** The same individual female face with a neutral expression (0% of intensity) and expressing happiness at five intensities (20%, 40%, 60%, 80%, 100% of expression). **C.** In each stimulation sequence, an individual neutral face (N) is presented through sinusoidal contrast modulation at a base rate of 6 Hz (1 cycle ≈ 167 ms). The same face expressing one emotion (E) is introduced at a lower rate of 1.2 Hz (every 5th cycle ≈ 833 ms between two expressions). Stimulus size varies randomly between 95% and 105% at every stimulation cycle. **D.** Each sequence starts with a 2.5-sec fade-in. Stimulation then lasts 100 s, followed by a 1-sec fade-out. Sequences are flanked by random pre- and post-stimulation intervals (0.5–1 s). After no expression-change during the fade-in, the intensity of the expression-changes increases every 20 s within the stimulation sequence from 20% to 100% of expression in 20% steps. Expression-change is set at 100% during the fade-out.

normal or corrected-to-normal visual acuity, and none reported any history of neurological or psychiatric illness. They provided written informed consent prior to beginning the experiment and were compensated for their participation. Testing was conducted in accordance with the Declaration of Helsinki.

Stimuli

The face stimuli were adapted from previous studies (e.g., Chambon et al., 2006; Durand et al., 2007; Leleu et al., 2016, 2015). A set of 24 color pictures from 4 individuals (2 females) displaying neutrality and 5 emotional expressions (anger, disgust, happiness, fear, sadness) in full-front view was used (Fig. 1A). Surprise was omitted due to its ambiguity (e.g., Kim et al., 2004). Each face was cropped into a medallion-shaped window discarding information about the background and body, and displayed on a mid-level grey (i.e., 128/255 in greyscale) background. For each emotional expression and individual face, various intensities were designed using linear continua of morphs combining neutrality and the expression (5 emotions × 4 individuals = 20 morphing continua) with Morpheus Photo Morpher 1.85 (Morpheus Software, USA). Five pictures were extracted from each continuum (from 20% to 100% of expression by steps of 20%). The final set of stimuli was thus composed of 104 pictures, 26 for each individual face (5 intensities × 5 emotions + neutrality). Pictures were set to a size of 6 × 4.8 cm (i.e., 6 × 4.8° of visual angle at a distance of 57 cm). Fig. 1B depicts an example of the five intensity steps used for an individual female face expressing happiness.

Procedure

The procedure was similar to the main experiment of Dzhelyova et al. (2017). Stimuli were presented on a 24-inch LED screen with a 60 Hz refresh rate and a resolution of 1920 × 1080 pixels. They were presented on a mid-level grey background (i.e., 128/255 in greyscale) through sinusoidal contrast modulation (from 0 to 100% of contrast) at a fast base rate of 6 Hz using custom Java software. At this rate, each stimulus lasts ≈167 ms (i.e., 1 s/6) with full contrast reached around 83 ms. In each stimulation sequence, the face of one individual with a neutral expression (N) was used as base stimuli. The same individual face expressing an emotion (E) was introduced every 5th stimulus, thus corresponding to an expression-change frequency of 6/5 = 1.2 Hz (i.e., ≈833 ms between two expressive faces). Schematically, the stimulation sequence was NNNNENN for ≈1.167 s of stimulation (Fig. 1C). To minimize expression-change detection based on low-level visual cues, picture size was randomly varied between 95% and 105% (i.e., 5.7 × 4.56° and 6.3 × 5.04° of visual angle respectively) at every stimulation cycle. Note that this FPVS paradigm is close to classical oddball paradigms used in ERP studies to measure "mismatch" responses (Näätänen et al., 1978) elicited by deviant (i.e., rare, 10–20% of occurrence) stimuli compared with standard (i.e., frequent, 80–90% of occurrence) stimuli. Mismatch responses in the visual domain following changes of facial expression have also been used as a signature of the detection of expression changes (e.g., Kimura et al., 2012; Stefanics et al., 2012; see Dzhelyova et al., 2017 for a detailed discussion). However, despite the interest of this approach, it suffers from most limitations of the standard ERP approach (see section 1). In particular, the long recording sessions needed to obtain sufficient SNR for analyzing mismatch responses prevent the evaluation of several facial emotions expressed at various intensities (see sections 4.1 and 4.3 for further discussion on classical and FPVS oddball paradigms).

Twenty-five conditions corresponding to the 5 emotions (anger, disgust, happiness, fear, sadness) × 5 intensities (20%, 40%, 60%, 80% and 100% of expression) were tested. One emotional expression was used throughout each sequence which started with a variable pre-stimulation interval of 0.5–1 s of a blank screen. It was followed by a 2.5-sec fade-in of increasing contrast modulation depth always presenting the neutral face (i.e., no expression-change). Then, the stimulation lasted 100 s. During the first 20 s, the emotional faces were expressed at 20% of intensity, followed by facial expressions at 40% of intensity during the next 20 s, followed by facial expressions at 60% of intensity during the next 20 s, and so on. In other words, during each 100-sec sequence, the intensity of the 1.2 Hz expression-change increased every 20 s in 20% steps from 20% to 100% of expression, as in an increasing sweep-VEP design (see Norcia et al., 2015). After the sequence, a fade-out of decreasing contrast modulation depth lasted 1 s with the expressive faces maintained at 100% of intensity, followed by a variable post-stimulation interval of 0.5–1 s of a blank screen. Fig. 1D illustrates the timeline of a sequence. Each emotion condition with increasing intensity steps was repeated 4 times (4 individual faces), resulting in 20 sequences. They were divided in 4 blocks of 5 sequences, each block presenting one sequence per condition. The presentation orders of blocks and sequences within blocks were randomized across participants.

After electrode-cap placement, participants were seated in a light- and sound-isolated cabin in front of the screen. Their head was held in place on a chinrest to be maintained at a distance of 57 cm from the screen and to reduce movements. An orthogonal behavioral task was designed to ensure that participants focused their attention on the screen. During each sequence, a fixation circle was always present at the center of the screen (located just below the eyes when the face stimuli were presented). Participants were asked to detect brief (200 ms) shape changes of the fixation circle (from circle to square) 10 random times within every sequence by pressing the space bar with both index fingers. A minimum interval of 6 s between two shape changes was introduced. When asked after the experiment, all participants reported having noticed expression-changes but none detected their periodicity.

EEG recording

During the experiment, electroencephalogram (EEG) was continuously recorded from a 64-channel BioSemi Active-Two amplifier system (BioSemi, The Netherlands) with Ag/AgCl electrodes located according to the 10–10 classification system. During recording, the Common Mode Sense (CMS) active electrode was used as reference and the Driven Right Leg (DRL) passive electrode was used as ground. Electrode offset was reduced between ±15 μV for each electrode. EEG was digitalized at a sampling rate of 1024 Hz.

EEG analysis

Preprocessing

All EEG analyses were performed using Letswave 5 (http://www.nocions.org/letswave5) running on Matlab 2012 (MathWorks, USA) and largely followed analyses steps described in Dzhelyova et al. (2017) and other studies (e.g., Retter and Rossion, 2016a). EEG data were first bandpass filtered at 0.1–100 Hz using a butterworth filter (4th order) and then downsampled to 256 Hz to reduce file size and processing time. EEG was cropped in 105-sec segments for each stimulation sequence according to 1 s before the fade-in and 0.5 s after the fade-out, thus resulting in 20 EEG segments (5 emotions × 4 individual faces). An Independent Component Analysis (ICA) was computed (e.g., Makeig et al., 1996) and components corresponding to eye blinks (recorded over Fp electrodes) and artifacts recorded over frontal and temporal electrodes were removed. Remaining noisy or artifact-ridden channels (i.e., containing activities exceeding ± 100 μV) in at least two trials for the same emotion were rebuilt using linear interpolation from the four nearest electrodes (mean number across participants: 0.61 channels, range: 0–4). EEG segments were then re-referenced to a common average reference.

Frequency-domain analysis

To reduce EEG activity non-phase-locked to the stimuli, the four preprocessed data segments obtained for each emotion were averaged in the time domain, thus resulting in a single 105-sec segment for each

emotion. These segments were further cropped in five shorter epochs corresponding to the five 20-sec periods for each intensity step within each sequence. Each epoch began at the onset of the first expressive face for one intensity condition (e.g., 0.667 s after the fade-in for 20% of intensity, 20.667 s after the fade-in for 40% of intensity, etc.) and lasted 20 s (i.e., exactly twenty-four 1.2 Hz cycles, 5120 time bins in total). A fast Fourier transform (FFT) was then applied to every epoch and amplitude spectra were extracted for all electrodes with a frequency resolution of $1/20 = 0.05$ Hz. Thanks to this high frequency resolution, 23 frequency bins were extracted between two frequencies of interest (e.g., between 1.2 and 2.4 Hz), thus allowing unambiguous identification of the base and expression-change responses and estimation of the noise level.

Group analysis. We first calculated the signal-to-noise ratio (SNR) for each emotion and intensity step on the FFT grand-averaged data across participants. SNR spectra were obtained by dividing the amplitude at one frequency bin by the mean noise amplitude estimated from the 20 surrounding frequency bins (10 on each side, excluding the immediately adjacent bins to safely exclude any contamination from the bin of interest in case of spectral leakage, and the 2 most extreme (minimum and maximum) bins to exclude potential outliers in the noise distribution; e.g., Dzhelyova et al., 2017; Liu-Shuang et al., 2014). SNR spectra were used for visualization and illustration purpose (Fig. 2), since responses at high frequencies in the EEG are generally of low amplitude but may have a high SNR.

We next determined how many harmonics (i.e., integer multiples; e.g., 6 Hz, 12 Hz, etc.) were significant for the base (6 Hz) response. After grand-averaging the FFT spectra (i.e., in amplitude, μV) across participants, electrodes, emotions and intensity steps, Z-scores were computed as the difference between the amplitude at one frequency bin and the mean noise amplitude (i.e., estimated from the 20 surrounding frequency bins, see above) divided by the standard deviation of the noise. Harmonics were considered significant until the Z-scores for two consecutive harmonics were no longer greater than 1.64 ($p < .05$, one-tailed, i.e., signal > noise). Significant harmonics were found until the 8th harmonic (48 Hz, i.e., harmonics were not considered after the 50 Hz response elicited by AC power). For the expression-change (1.2 Hz) response expected to be nearly absent at 20% of expression intensity and to increase with increasing intensity steps, the number of significant harmonics

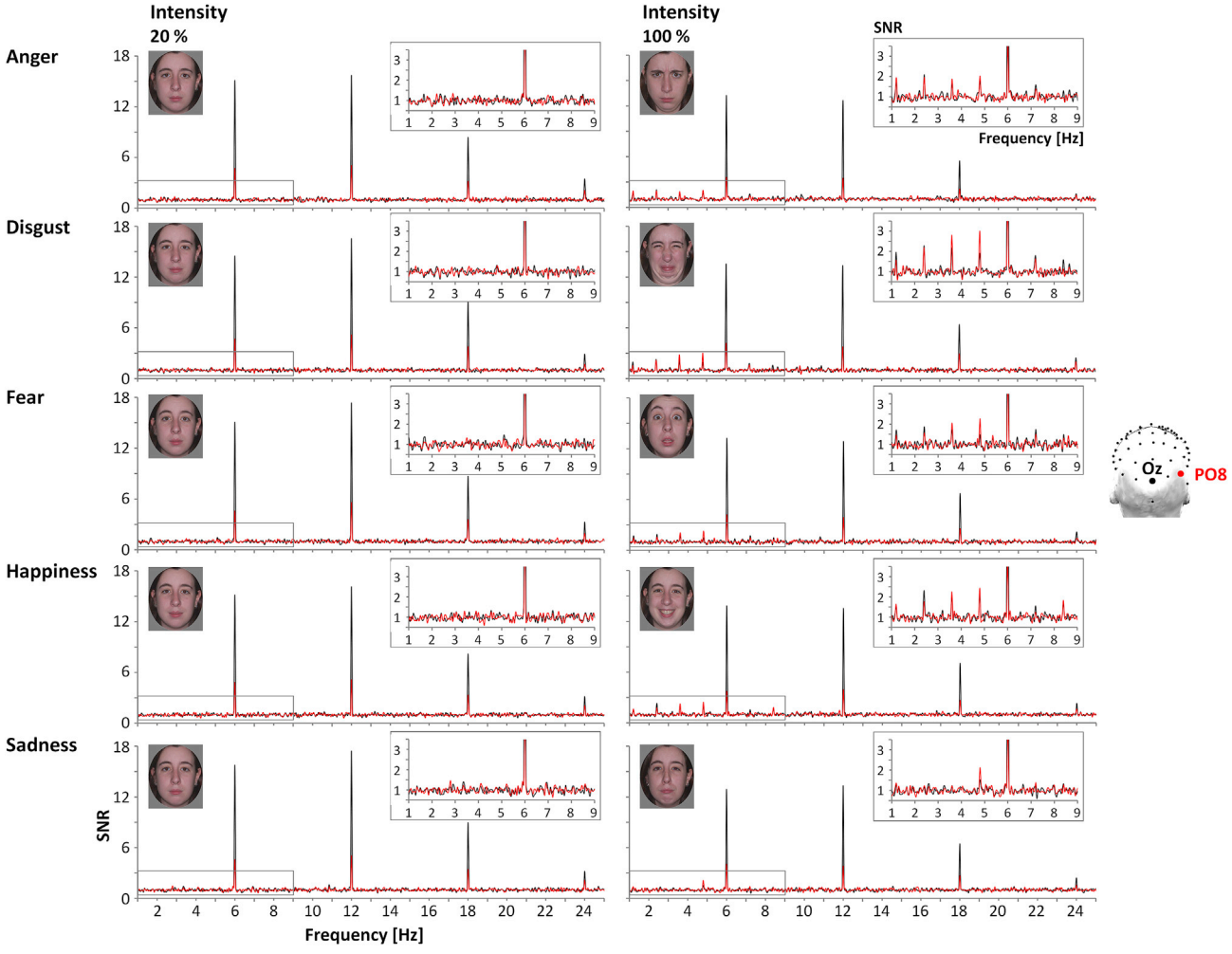


Fig. 2. Grand-averaged FFT signal-to-noise ratio (SNR) spectra. SNR calculated on the grand-averaged FFT amplitude spectra for the different emotions expressed at 20% (left) and 100% (right) of intensity (displayed from 1 to 25 Hz) over the medial occipital channel Oz (black) and the right occipito-temporal channel PO8 (red). For each condition, the inset zooms in the SNR elicited from 1 to 9 Hz. Responses at the expression-change rate (1.2 Hz) and its harmonics (i.e., 2.4 Hz, 3.6 Hz, etc.) are only visible for facial expressions at 100% of intensity, with greater SNR over PO8 than Oz, reflecting the sensitivity to high intensity changes of facial expressions. The highest and the lowest expression-change responses are found for disgust and sadness respectively. In contrast, high SNR responses are clearly visible at the base rate (6 Hz) and its harmonics (i.e., 12 Hz, 18 Hz, etc.) for all conditions, mainly over Oz, reflecting the general sensitivity to all visual cues rapidly changing at this frequency. These base responses appear lower at 100% than 20% of intensity.

should vary as a function of expression intensity. Hence, according to Dzhelyova et al. (2017) who found significant harmonics up to 16.46 Hz with similar stimulation parameters, harmonics were considered until the 14th harmonic (i.e., 16.8 Hz).

Z-scores were then calculated on FFT data summed until the 8th harmonic for the base response and until the 14th harmonic (excluding the two harmonics corresponding to the base rate; i.e., 6 and 12 Hz) for the expression-change response. Summed amplitudes across harmonics were used to quantify the overall response in the frequency-domain (Retter and Rossion, 2016a). In average across emotions and intensity steps, all electrodes over the scalp reached significance for the base response (all $Zs > 9.75$, all $ps < .001$, greatest $Z = 201.84$ for Oz) while 29 electrodes reached significance for the expression-change response (greatest $Z = 4.42$ for PO8). Considering these significant channels, *Z*-scores were calculated for the expression-change response on summed amplitudes averaged across emotions in order to assess the lowest intensity step eliciting a significant response.

We then applied a data-driven approach to determine different regions-of-interest (ROIs) to include in statistical analyses for both base and expression-change responses. For the base response, ROIs were determined using the grand-averaged data pooled across all conditions (i.e., 5 emotions × 5 intensities) since its topography should not differ among the emotions (Dzhelyova et al., 2017). In contrast, for the expression-change response, ROIs were determined separately for each emotion using the grand-averaged data pooled only across intensities, since different topographies were previously observed depending on the emotion expressed (Dzhelyova et al., 2017). Hence, we excluded differences between emotions that would be only explained by topographical differences if we had considered similar ROIs, especially for their respective sensitivity to expression intensity. To scale differences between electrodes on the global magnitude of the response across significant channels previously identified and thus define more precise ROIs, the summed amplitudes at each channel were first normalized by dividing by the square root of the sum of squared amplitudes of these channels (McCarthy and Wood, 1985). This procedure allows identifying the electrodes over which the response is largest irrespective of its global power. Then, *Z*-scores were calculated on normalized summed amplitudes and only channels with significant responses were included in three ROIs: left and right posterior sites (LH and RH respectively), and medial occipital sites (MO). The electrodes included in each ROI thus differed according to the response (base *vs.* expression-change) and the emotion expressed when considering the expression-change response (Supplementary Fig. S1).

Finally, both responses were quantified in a single value expressed in microvolt (μV) for statistical analyses. A baseline-correction was first applied to FFT amplitude spectra by subtracting the mean amplitude of the noise (i.e., estimated from the 20 surrounding bins, see above). Then, these baseline-corrected amplitudes (BCA) were summed until the 8th harmonic (i.e., 48 Hz) for the base response, and until the 14th harmonic (i.e., 16.8 Hz) for the expression-change response, excluding the two harmonics corresponding to the base rate (i.e., 6 and 12 Hz). Summed BCA were calculated for every condition and participant. Grand-averages of summed BCA were computed for data visualization.

Repeated-measures ANOVAs were then run on individual summed BCA data for both base and expression-change responses with *Emotion* (anger, disgust, happiness, fear, sadness), *Intensity* (20%, 40%, 60%, 80%, 100%) and *ROI* (MO, LH, RH) as within-subject factors. Mauchly's test for sphericity violation was performed and Greenhouse-Geisser correction for degrees of freedom was applied whenever sphericity was violated. For significant effects, post-hoc comparisons were conducted using Tukey's HSD test. Since the tuning functions to expression-changes depending on intensity are the main focus of this study, polynomial contrasts were calculated to estimate the relationship between the amplitude of the expression-change response and expression intensity.

Individual analysis. Given the high SNR of the technique, we also measured how much expression intensity was sufficient to elicit a significant expression-change response in each individual participant. These neural detection thresholds were determined using the same procedure as for group-level analysis (see section 2.5.2.1). To evaluate the lateralization of the effect of expression intensity in each participant, individual expression-change responses at 20% of intensity were subtracted from those at 100% of intensity and summed BCA over all relevant lateral electrodes identified in the group analysis were averaged for each hemisphere. A lateralization index was then computed [(RH − LH)/(RH + LH)] with positive and negative values revealing right- and left-sided responses respectively.

Time-domain analysis

In order to examine the time-course of the expression-change response depending on expression intensity, a time-domain analysis was performed on preprocessed data. A 20 Hz low-pass FFT filter was first applied with a 1 Hz cutoff. To selectively remove all the periodic visual responses time-locked to the base stimulation, an FFT multinotch narrowband filter with a 0.5 Hz width was applied to the 6 Hz base rate and its harmonics (e.g., the 5.75–6.25 Hz frequency band was selectively removed for the base frequency, the 11.75–12.25 Hz band for its second harmonic, etc.) After removing the data recorded during the fade-out, EEG segments were cropped in 834 ms-long (214 time bins) epochs according to 167 ms before and 667 ms after every expression-change onset (i.e., corresponding to a NENNN sequence), resulting in 96 epochs per condition (20 s of 1.2 Hz cycles × 4 individual faces). These epochs were baseline corrected by subtracting the mean amplitude during the 167 ms before the expression-change onset. Epochs containing amplitudes greater than ±100 μV at the relevant lateral electrodes identified in the frequency-domain analysis were discarded (mean number across participants and conditions: 0.54 epochs, range: 0–11). Significant differences in epoch rejection were neither found according to *Emotion*, $F(1.4, 24.2) = 1.05$, $\varepsilon = .36$, $p = .34$, *Intensity*, $F(1.8, 29.8) = 1.17$, $\varepsilon = .44$, $p = .32$, nor their interaction, $F < 1$. The remaining epochs were then averaged for each participant and grand-averaged for visualization purpose.

Since the visual processes common to all stimuli (i.e., elicited at each stimulation cycle) are captured at 6 Hz and harmonics, 6 Hz filtered-out data reflect direct differential activities specific to the detection of an expression-change (Dzhelyova et al., 2017). In other words, they represent the expression-change response isolated at 1.2 Hz and harmonics. Hence, we averaged all electrodes showing a significant expression-change response in the frequency-domain and determined the time-windows when the waveform deflections consistently differ from zero to define the different expression-change specific components. *T*-tests against zero were computed at each time bin on the average across conditions. To reduce the risk of false positives, a criterion of 8 consecutive significant bins (≈31 ms) between 0 and 667 ms after expression-change onset was used to determine the time-course of different components.

As different channels might contribute to each component depending on the emotion expressed (Dzhelyova et al., 2017), a data-driven approach was then used to determine the electrodes pooled together in different ROIs for each component and each emotion. Mean amplitudes were calculated within each time-window for each emotion averaged across intensities and for each electrode previously considered. As for the frequency-domain analysis (see section 2.5.2.1.), normalization (McCarthy and Wood, 1985) was applied to isolate the electrodes showing the largest response and their significance was determined using *T*-tests against zero. Significant channels were finally averaged in LH and RH ROIs for each emotion and time-window (Supplementary Fig. S2).

Finally, mean amplitudes (before normalization) within each time-window were submitted to repeated-measures ANOVAs considering three within-subject factors: *Emotion* (anger, disgust, happiness, fear, sadness), *Intensity* (20%, 40%, 60%, 80%, 100%) and *Hemisphere* (LH,

RH). Mauchly's test for sphericity violation was used and Greenhouse-Geisser correction was applied whenever necessary. For significant effects, post-hoc comparisons were conducted using Tukey's HSD test. *T*-tests against zero were also performed to assess response significance depending on intensity. Polynomial contrasts were finally calculated to determine the tuning functions to expression-changes as a function of intensity.

Results

Behavior

The shape-change detection task was well performed, with accuracy near ceiling ($M = 96.1 \pm 1.1$ (*SEM*) %) and mean RTs of $M = 453 \pm 12$ ms. This indicates that participants paid full attention to the screen during the periodic stimulation.

EEG

Frequency-domain

In accordance with the main experiment of Dzhelyova et al. (2017), SNR calculated on the FFT amplitude spectra (Fig. 2) show that the brief periodic changes from neutrality to emotional expressions at 100% of intensity elicit clear brain responses at the 1.2 Hz expression-change rate and harmonics for all emotions (i.e., SNR between 1.5 and 3). Interestingly, however, no identifiable responses are observed for emotions expressed at 20% of intensity (i.e., SNR ≈ 1). In comparison, the 6 Hz base rate elicits synchronized periodic EEG activities of high amplitudes, with the signal around 6 to 18 times larger than the noise for all facial expressions and all harmonics displayed in Fig. 2 (i.e., until 24 Hz). Note that these responses are higher for emotions expressed at 20% *vs.* 100% of intensity (i.e., during the first than the last 20 s of the sweep stimulation sequence).

Automatic detection of brief facial expression-changes quantified in the brain from the sweep increase of expression intensity. For all emotions pooled together, visual inspection of the topographical maps of summed BCA suggests that the expression-change response appears as a function of the sweep increase of expression intensity mainly in the right occipito-temporal regions (Fig. 3). This was confirmed by a significant main effect of *Intensity*, $F(1.6, 27.8) = 34.47$, $\varepsilon = .41$, $p < .001$, $\eta_p^2 = .67$, qualified by an *ROI* × *Intensity* interaction, $F(4.5, 76.5) = 4.12$, $\varepsilon = .56$, $p = .003$, $\eta_p^2 = .19$. Brief expression-changes were practically not detected at 20% of intensity ($M = .08 \pm .03$ μV), while they elicited increased amplitudes as the intensity of the expression increased (from $M = .24 \pm .04$ μV at 40% of intensity to $1.05 \pm .15$ μV at 100% of intensity) for an overall signal increase of $M = .97 \pm .15$ μV. This amplitude increase was larger in RH ($M = 1.07 \pm .14$ μV) and MO ($M = 1.03 \pm .17$ μV) than LH ($M = .81 \pm .16$ μV) ROIs. Based on a criterion considering significant responses over one channel as soon as *Z*-scores are above 1.64 ($p < .05$, one-tailed, i.e., signal > noise) for two consecutive intensities, the lowest intensity inducing a discrimination of facial expression from neutrality was 40% with the largest significant response over PO7 ($Z = 2.30$; Fig. 3). Significant responses were further observed from 60 to 100% of expression intensity with the greatest response at 100% of intensity found over the right occipito-temporal channel P8 ($Z = 11.62$).

The tuning functions to expression-changes at the five intensities were explored through polynomial contrasts. They revealed that the relationship between expression intensity and the strength of the expression-change response was a mixture of a highly significant linear component, $F(1, 17) = 44.77$, $p < .001$, $\eta_p^2 = .72$, and a quadratic component, $F(1, 17) = 6.22$, $p = .023$, $\eta_p^2 = .27$. No other contrasts were significant, all $Fs < 1.53$, all $ps > .23$. Thus, the expression-change response increased predominantly, but not completely, as a linear function of expression intensity (16.9% of the overall signal increase occurred between 20 and 40% of intensity, 20.0% occurred between 40 and 60% of intensity, 35.9% between 60 and 80%, and 27.2% between 80 and 100% of intensity). The more abrupt increase beyond 60% of intensity may likely explain the quadratic component.

In addition, to ensure that the significant amplitude increase of the expression-change response as a function of intensity is not only elicited by increasing physical changes between neutral and expressive faces, we calculated a physical dissimilarity index for each emotion and intensity step. Pictures from neutral and expressive faces at each intensity step were first converted to greyscale images for each emotion and individual face. Pixel-wise correlations between neutral and expressive faces were then computed and averaged across individual faces. By subtracting these correlations from 1 and multiplying by 100, a physical dissimilarity index was obtained for each emotion (i.e., 25.4, 26.4, 18.9, 13.7 and 13.4 respectively for anger, disgust, fear, happiness and sadness at 100% of intensity) with five linear steps reflecting the linear morphs created for simulating the increase of expression intensity (e.g., 5.1, 10.2, 15.2, 20.3 and 25.4 for anger). Finally, summed BCA were normalized for each participant, emotion and intensity by dividing BCA values by the corresponding physical dissimilarity indexes. The main effect of *Intensity* was still significant after this procedure, $F(2.1, 35.8) = 6.83$, $\varepsilon = .53$, $p = .003$, $\eta_p^2 = .29$, thus confirming that increasing physical changes between neutral and expressive faces as a function of intensity does not fully explain amplitude increase of the brain response to brief changes of expression.

For the modulation of the base response along the sweep stimulation sequence, an opposite pattern was found. Topographical maps of summed BCA (Fig. 3) show a medial occipital response centered on channel Oz already observed when expression-changes are at 20% of intensity (i.e., during the first 20 s of the sequence) and decreasing for expressions at higher intensities (i.e., at each following 20 s steps in the sweep sequence). The main effect of *Intensity*, $F(1.5, 26.2) = 41.52$, $\varepsilon = .38$, $p < .001$, $\eta_p^2 = .71$, confirmed that the amplitude of the base response was greater during the first 20 s of the stimulation sequence ($M = 3.13 \pm .34$ μV) and further decreased at each following 20 s step (from $M = 2.61 \pm .29$ μV at 40% to $M = 2.43 \pm .27$ μV at 100% of intensity). In other words, the signal elicited at the base rate was reduced by 16.7% after the first 20 s of stimulation, for an overall decrease of 22.3% along the entire stimulation sequence. Thus, 74.9% of the overall decrease occurred during the first 20 s. Accordingly, the *Intensity* effect was mainly driven by a significant reduction in amplitude between 20% and all other intensities (all $ps < .001$). A slight difference appeared between 40 and 100% of intensity ($p = .042$) but no other differences were significant (all $ps > .075$).

The medial occipital topography of the base response was supported by the main effect of *ROI*, $F(2, 34) = 16.81$, $p < .001$, $\eta_p^2 = .50$, with greater summed BCA for MO ($M = 3.53 \pm .42$ μV) compared to RH ($M = 2.36 \pm .30$ μV, $p < .001$) and LH ($M = 2.06 \pm .25$ μV, $p < .001$) sites. The two latter ROIs did not differ significantly ($p = .52$). Activities over MO thus account for 44.4% of the signal recorded over the three posterior ROI (29.7 and 25.9% respectively for RH and LH). In addition, a significant *ROI* × *Intensity* interaction, $F(3.6, 61.6) = 4.28$, $\varepsilon = .45$, $p = .005$, $\eta_p^2 = .20$, further showed that the amplitude decrease of the base response was greater over MO (from $M = 4.10 \pm .48$ μV at 20% to $M = 3.28 \pm .40$ μV at 100% of intensity) than RH (from $M = 2.86 \pm .35$ μV to $M = 2.13 \pm .28$ μV) and LH (from $M = 2.42 \pm .29$ μV to $M = 1.88 \pm .23$ μV) ROIs. However, these differences in absolute signal reduction are quite equivalent relative to the strength of the response over each ROI, with a signal decrease of 20.0 and 22.3% for MO and LH sites respectively, and a slightly larger decrease of 25.5% for RH sites.

Distinct patterns of expression-change responses between emotions irrespective of expression intensity. As in the previous study of Dzhelyova et al. (2017), the expression-change response presents with distinct topographies and magnitudes, according to the category of emotion expressed (Fig. 4).

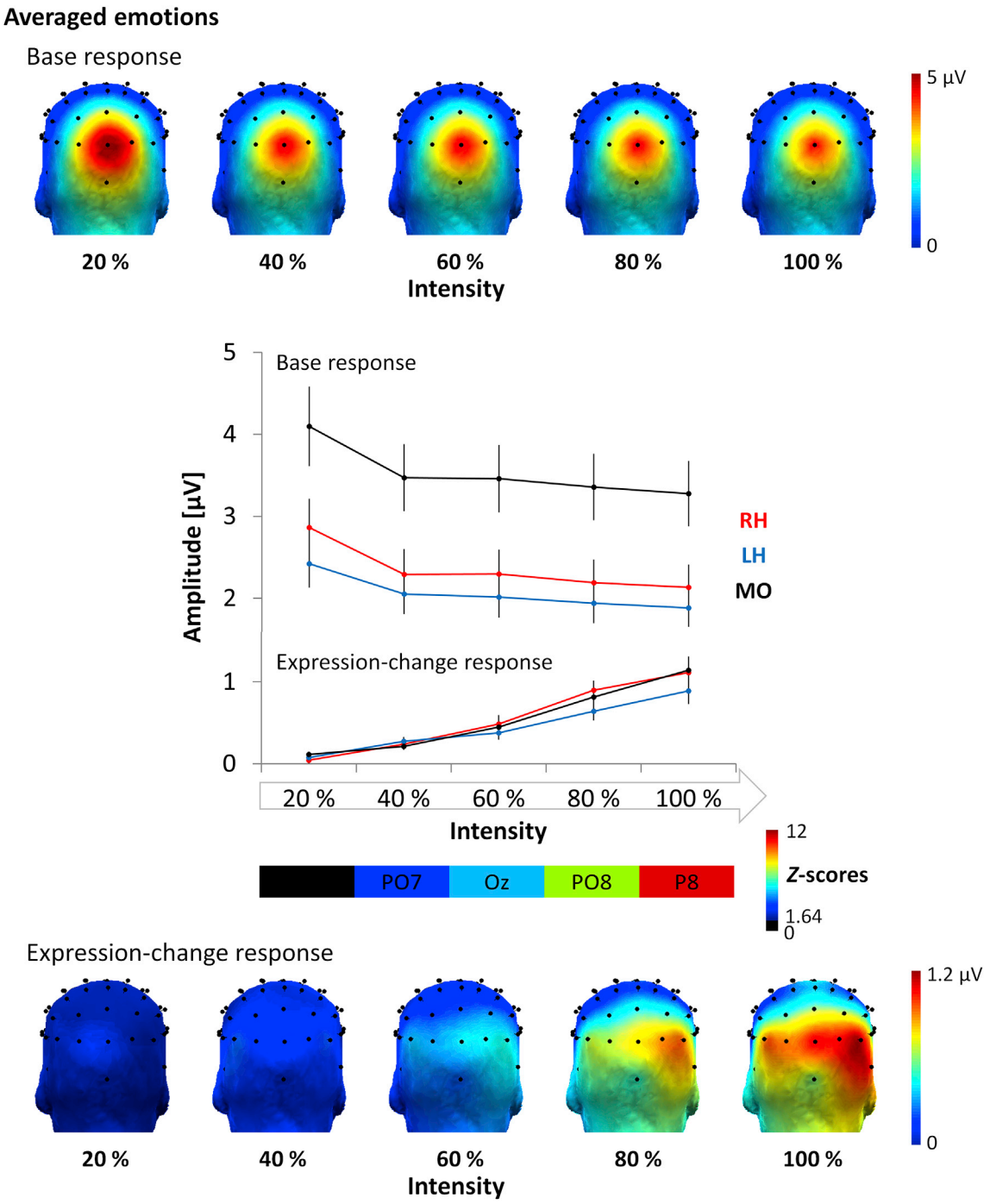


Fig. 3. Dissociable effect of the sweep increase of expression intensity between the base and the expression-change responses. 3D-topographical maps (posterior view) of summed baseline-corrected amplitudes (BCA) for each expression intensity averaged across emotions, and for the base (top) and the expression-change (bottom) responses. Center: the same data displayed for medial occipital (MO), left (LH) and right (RH) occipito-parietal and occipito-temporal ROIs (channels pooled in each ROI depend on the response and on the emotion for the expression-change response; Supplementary Fig. S1). Bars represent standard errors of the mean. For the base response, large amplitudes are clearly visible over medial occipital sites centered on channel Oz. Note the decrease in amplitude mainly observed between 20% and 40% of intensity (i.e., during the first 20 s of the sweep stimulation). In contrast, the expression-change response is nearly absent when emotions are expressed at 20% of intensity and it progressively increases as a function of expression intensity, mainly over right occipito-temporal regions. A color-coded bar above the maps for the expression-change response depicts the strength of the largest *Z*-score at each intensity step and its corresponding channel, revealing that the first detection of expression-change from neutrality appears at 40% of intensity.

Visual inspection of topographical maps of normalized (McCarthy and Wood, 1985) summed BCA calculated across intensities first highlight scalp topographical differences between emotions (Fig. 4A). They suggest that happiness is characterized by an occipito-parietal topography while other emotions are characterized by a more ventral occipito-temporal response. Fear and sadness give rise to more restricted distributions, fear being the only emotion whose activities are more equally distributed across hemispheres. These different spatial patterns were also revealed by the selection of electrodes included in each ROI (Supplementary Fig. S1).

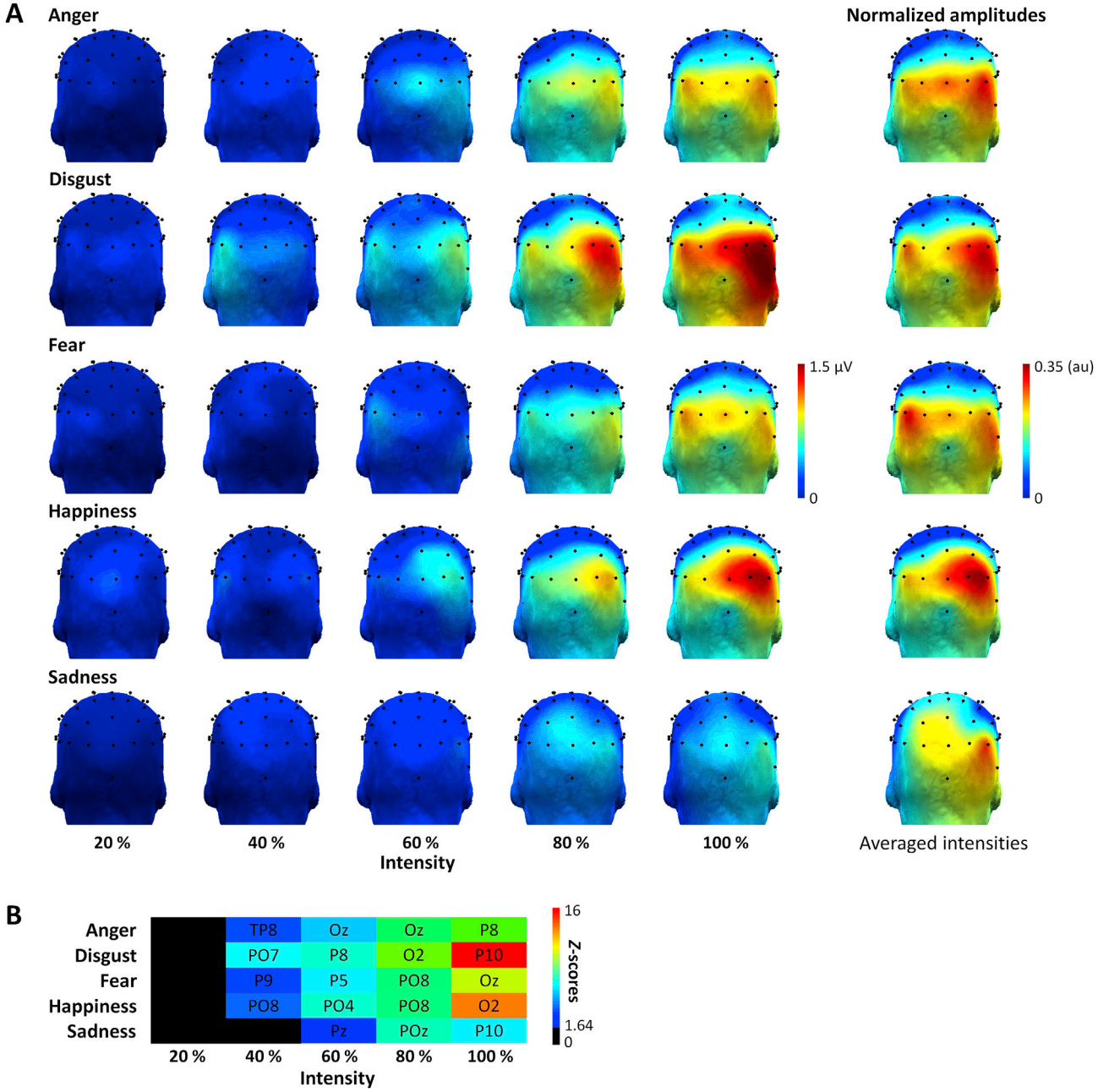


Fig. 4. Brain responses to brief facial expression-changes of various emotions during the sweep increase of expression intensity. A. 3D-topographical maps (posterior view) of summed baseline-corrected amplitudes (BCA) for each emotion depending on expression intensity, and maps of normalized summed BCA averaged across intensities for each emotion. Note the increase in amplitude with increasing intensity steps, mainly over channels in the right hemisphere, and greater for disgust and happiness while lower for sadness. On the normalized maps, dorsal *vs.* ventral topographies for happiness *vs.* the other emotions are notably visible (au: arbitrary unit). **B.** Color-coded table representing the strength of the largest *Z*-score at each intensity step and its corresponding channel for each emotion, revealing the lowest intensity that elicits a significant detection of expression-change from neutrality. The first response appears at 40% of intensity for all emotions except sadness (60%).

Inspection of the maps further suggests larger expression-change responses for disgust and happiness and lower for sadness (Fig. 4A). Accordingly, the main effect of *Emotion*, $F(4, 68) = 13.52$, $p < .001$, $\eta_p^2 = .44$, revealed larger summed BCA for disgust ($M = .73 \pm .10\,\mu V$) compared to all other emotions (anger: $M = .52 \pm .08\,\mu V$, $p = .009$, fear: $M = .49 \pm .06\,\mu V$, $p = .002$, sadness: $M = .29 \pm .06\,\mu V$, $p < .001$) except happiness ($M = .56 \pm .08\,\mu V$, $p = .063$). Amplitude was significantly lower for sadness than for all other emotions (all $ps < .015$). No other differences between emotions were observed (all $ps > .71$).

Similarly to the *Intensity* effect detailed in section 3.2.1.1., we determined whether the response may have been driven by the physical differences between neutral and expressive faces. After normalization of the data using the physical dissimilarity indexes previously calculated, the *Emotion* effect was still significant, $F(2.5, 42.6) = 10.94$, $\varepsilon = .63$, $p < .001$, $\eta_p^2 = .39$. Hence, this analysis suggests that despite a presumable influence of low-level physical differences between neutrality and each expression, the pattern of amplitude is driven by other high-level perceptual processes.

Interestingly, despite the above evidence for different patterns of expression-change responses between emotions, the interaction between *Emotion* and *Intensity* was not significant, $F(7.1, 120.1) = 1.81$, $\varepsilon = .44$, $p = .091$, suggesting that the tuning functions for automatic detection of

brief expression-changes are not different across all tested emotions. Accordingly, using the criterion considering significant responses as soon as *Z*-scores were above 1.64 for two consecutive intensities, no expression-change elicited significant responses at 20% of intensity. The expression-change responses appeared for all emotions expressed at 40% of intensity (from $Z = 1.84$ over P9 for fear to $Z = 4.71$ over PO7 for disgust) except sadness. Significant responses were further observed for all emotions from 60 to 100% of intensity (Fig. 4B).

In comparison, for the base response, there was no main effect of *Emotion*, $F(4, 68) = 1.09$, $p = .37$, or of all other effects involving the *Emotion* factor, all Fs < 1.28, all ps $> .14$, showing the absence of significant differences between emotions (range: $M = 2.59$–$2.70\,\mu V$). Since EEG activities recorded at the base rate and harmonics reflect the processes elicited by low- and high-level visual information that changes at 6 Hz, these observations confirm that these processes are involved for all emotions and do not differ between them.

Individual neural thresholds for automatic detection of brief facial expression changes. Thanks to the high sensitivity of the FPVS-EEG approach, we determined neural thresholds for detecting brief changes of facial expression in each individual participant (Fig. 5A). Given that the effect of *Intensity* did not interact with the emotion expressed at the group level, individual responses were averaged across emotions and their significance was estimated using the same criterion as for the group-level analysis (see previous sections). Significant expression-change responses were already observed at 20% of intensity for 5 out of the 18 participants (Fig. 5A.). The number of participants presenting a significant response rapidly increased with intensity steps with 12 participants at 40% of expression intensity and all participants at 60%, 80% and 100% of intensity. In other words, individual detection thresholds are estimated from our EEG data with 5 participants already detecting expression-changes at 20% of expression intensity, 7 participants at 40% of expression intensity, and the 6 last participants at 60% of intensity.

In line with the finding from the grand-averaged data showing that the increase of the expression-change response as a function of expression intensity is larger over right posterior brain regions, the largest individual responses progressively appear over right posterior channels during the sweep increase of intensity and are prevalent over these channels when emotions are expressed at 80% and 100% of intensity (Fig. 5A). Thirteen out of the 18 participants presented their largest response at 80% and/or 100% of intensity over P10, PO8, PO4, or O2. Fig. 5B depicts the topographical maps of summed BCA obtained for the difference between 100% and 20% of expression intensity and illustrates this right hemisphere advantage across individual participants. The lateralization index quantifying the additional proportion of signal recorded over the dominant hemisphere (i.e., positive vs. negative for right-vs. left-sided responses) for these differential responses between 100% and 20% of expression intensity showed a RH advantage of +15.3% at the group level. Calculated for each individual response, it revealed a right-sided asymmetry in 15 out the 18 participants (from +0.9% for S05 to +259.2% for S07; remaining 3 participants: S15, S16, S18, from −10.2% for S15 to −20.0% for S18).

Time-domain: temporal dynamics of the brain tuning to facial expression changes

By filtering out the signal recorded at the base rate and its harmonics from the EEG data, the responses elicited by the periodic modulation of all visual cues rapidly changing at 6 Hz were selectively removed, thus providing direct differential expression-change specific activities in the time-domain (Dzhelyova et al., 2017). A triphasic response reflecting the discrimination of an emotional expression from neutrality was identified until approximately 500 ms after expression-change onset (Fig. 6). The three components sequentially peaked at 140, 240 and 350 ms post expression-change (time-windows: 75–160 ms, 219–277 ms, 334–406 ms) and were respectively positive, negative and positive.

For the first positive response peaking at 140 ms after expression-change onset mainly over dorsal sites, a significant main effect of *Intensity*, $F(4, 68) = 13.54$, $p < .001$, $\eta_p^2 = .44$, showed increased amplitudes with increasing intensity (from $M = 0.05 \pm 0.05\,\mu V$ at 20% to $0.38 \pm 0.06\,\mu V$ at 100% of intensity) for an overall signal increase of $M = .33 \pm .08\,\mu V$. A marginally significant response emerged at 40% of intensity, $T(17) = 1.99$, $p = .062$, and highly significant responses were then found for emotions expressed at 60% of intensity and more (from 60 to 100%: all Ts > 5.54, all ps $< .001$, 20%: $T(17) = 1.13$, $p = .28$). Polynomial contrasts revealed a linear relationship between expression intensity and the amplitude of the first component, $F(1, 17) = 29.63$, $p < .001$, $\eta_p^2 = .64$. No other significant contrasts were found, all Fs < 1.85, all ps $> .19$. Neither the main effect of *Emotion*, $F(4, 68) = 1.04$, $p = .39$, nor the *Emotion* × *Intensity* interaction, $F < 1$, were significant, indicating no specific sensitivities to expression intensity across emotions.

For the negative component peaking at 240 ms mainly over ventral channels, the main effect of *Intensity*, $F(2.4, 40.6) = 4.21$, $\varepsilon = .60$, $p = .017$, $\eta_p^2 = .20$, revealed amplitude increase with increasing expression intensity (from $M = -.19 \pm .08\,\mu V$ at 20% to $-.64 \pm .16\,\mu V$ at 100% of intensity, overall signal increase: $M = -.45 \pm .15\,\mu V$). The component was sensitive to low-intensity expression-change with a significant response already found at 20% of intensity ($T(17) = 2.27$, $p = .037$, other intensities of expression: all Ts > 3.71, all ps $< .002$). Polynomial contrasts revealed a significant linear relationship between expression intensity and the strength of the second component, $F(1, 17) = 7.24$, $p = .015$, $\eta_p^2 = .30$. No other contrasts were significant, all Fs < 1.47, all ps $> .24$. Neither the main effect of *Emotion*, $F(2.6, 44.3) = 1.71$, $\varepsilon = .65$, $p = .18$, nor its interaction with *Intensity*, $F(7, 119.5) = 1.13$, $\varepsilon = .44$, $p = .35$, were significant. Hence, as for the first component, no amplitude differences between emotions or their modulation by expression intensity were evident.

Finally, for the third component occurring 350 ms after expression-change onset and encompassing a large posterior topography, the main effect of *Emotion* was significant, $F(4, 68) = 4.29$, $p = .004$, $\eta_p^2 = .20$, with greater amplitudes for anger ($M = .44 \pm .10\,\mu V$) and disgust ($M = .55 \pm .13\,\mu V$) than for sadness ($M = .11 \pm .12\,\mu V$, all ps $< .038$). No other differences between emotions were found (all ps $> .13$, fear: $M = .28 \pm .11\,\mu V$, happiness: $M = .37 \pm .09\,\mu V$). All emotions elicited significant responses (all Ts > 2.49, all ps $< .024$) except sadness, $T(17) = .89$, $p = .39$. In addition, a significant main effect of *Intensity*, $F(2, 33.8) = 24.95$, $\varepsilon = .50$, $p < .001$, $\eta_p^2 = .59$, indicated increased amplitudes with increased intensities of expression (from $M = -.02 \pm .07\,\mu V$ at 20% to $.88 \pm .17\,\mu V$ at 100% of intensity) for an overall signal increase of $M = .89 \pm .18\,\mu V$. The component was only sensitive to moderate to high-intensity expression-changes since significant responses were found for emotions expressed at 60%, $T(17) = 2.13$, $p = .049$, 80%, $T(17) = 5.27$, $p < .001$, and 100% of intensity, $T(17) = 5.15$, $p < .001$, but not 20% and 40% (all Ts $< .37$, all ps $> .71$). Accordingly, and interestingly, only 25.2% of the overall signal increase occurred between 20 and 60% of intensity while the remaining 74.8% occurred between 60 and 100% of intensity (from 60 to 80%: 54.9%, from 80 to 100%: 19.9%). Hence, polynomial contrasts revealed several relationships between expression intensity and the amplitude of the component, both linear, $F(1, 17) = 33.49$, $p < .001$, $\eta_p^2 = .66$, quadratic, $F(1, 17) = 7.08$, $p = .016$, $\eta_p^2 = .29$, and cubic, $F(1, 17) = 11.81$, $p = .003$, $\eta_p^2 = .41$, contrasts being significant. The quartic contrast was not significant, $F(1, 17) = 2.03$, $p = .17$. The *Intensity* effect was qualified by an *Emotion* × *Intensity* interaction but it did not survive Greenhouse-Geisser correction, $F(6.6, 111.4) = 1.72$, $\varepsilon = .41$, $p = .12$, $\eta_p^2 = .09$, before correction $p = .044$. This marginally significant interaction showed lower amplitude increase with intensity steps for sadness ($M = .35 \pm .19\,\mu V$) compared to the other emotions ($M = 1.03 \pm .19\,\mu V$). As a result, while amplitudes for 80 and 100% of intensity were highly significant for all other emotions (all Ts > 2.93, all ps $< .01$), they only tend to reach significance for sadness (all Ts < 2.01, all ps $> .061$).

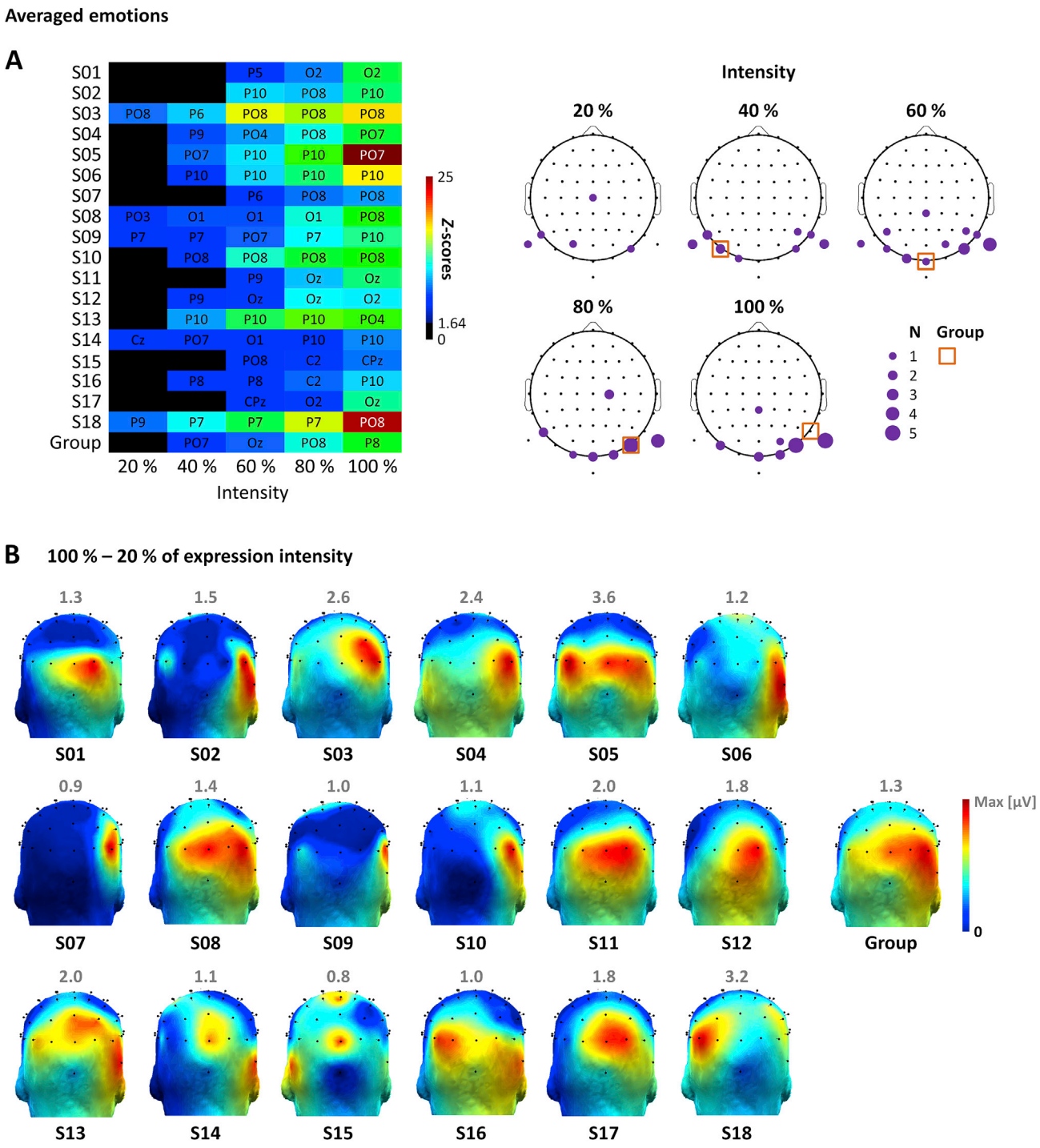


Fig. 5. Individual thresholds for automatic detection of brief changes of facial expression. A. Left: color-coded table depicting the strength of the greatest *Z*-scores at each intensity step averaged across emotions and its corresponding channel for each individual participant and for the group. Right: 2D-topographical maps of the same data depicting the density of significant individual responses over each channel. Bubble size at every single channel reflects the number of participants presenting their greatest Z-score here. The orange square points to the electrode with the greatest Z-score for the group. Note the progressive consistency of the response over right posterior regions as a function of expression intensity. **B.** 3D-topographical maps (posterior view) of summed baseline-corrected amplitudes (BCA) of expression-change responses at 20% of intensity subtracted from those at 100% of intensity averaged across emotions for each participant and for the group. The grey numbers above each map indicate the magnitude of individual color-scales.

Discussion

Our observations extend from recent findings quantifying the ability of the human brain to automatically discriminate brief changes from a neutral expression to an expressive face (Dzhelyova et al., 2017) with a different stimulus set and two additional facial expression changes (anger, sadness). Here, we introduced a sweep increase of intensity to changes of expression showing robust facial expression detection already at 40% of intensity for most of the basic emotions (anger, disgust, fear and happiness) and at 60% of intensity for sadness. The

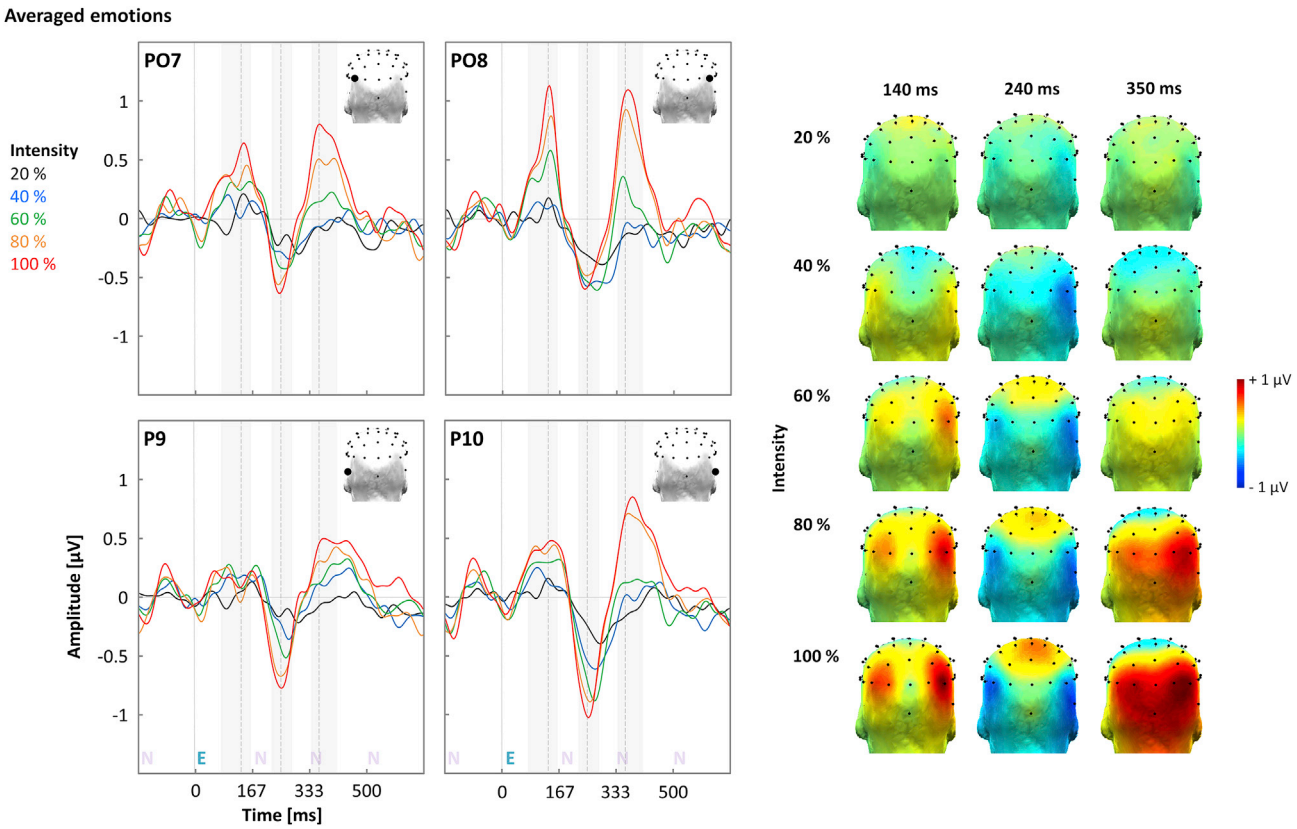


Fig. 6. Time-course of the responses to brief facial expression-changes of increasing intensity. Left: 6 Hz filtered-out grand-averaged EEG data recorded between 167 ms before and 667 ms after the expression-change onset over channels PO7/8 and P9/10 for each expression intensity averaged across emotions. Three deflections peaking at 140, 240 and 350 ms and of respective positive, negative and positive amplitudes are elicited by the discrimination of an expression from neutrality and are clearly visible until 500 ms after expression-change onset. The grey periods represent the time-windows and the dashed lines the latency for the maps. Stimulus-onsets (N: neutral, E: expression) are represented under the waveforms and attenuated representation of the neutral face stimuli symbolizes that the activities recorded at the base rate were excluded. Right: 3-D topographical maps (posterior view) of each component showing linear increase of amplitudes with increasing intensity for the two first components and a more abrupt increase for the third component. Distinct topographies are also visible with dorsal *vs.* ventral activities respectively for the first and second components, the third component eliciting a large posterior topography.

expression-change response increases in amplitude with increasing intensity steps for each emotion tested, predominantly over right occipito-temporal sites, thus constituting the first probing electrophysiological evidence of how various facial expressions are automatically detected as a function of intensity. In addition, this neural marker was quantified at the individual level to provide a direct measure of how much expression intensity is sufficient to elicit a significant detection of facial expression in each individual brain. Interestingly, the brain sensitivity to brief changes of facial expression over time reveals two-step tuning functions along three components: the two first components show linear detection likely related to low-level image changes and visual coding of facial movements, followed by categorical perception of emotional expressions from about 300 ms at the level of the third component.

A robust neural response to brief changes of facial expression

Thanks to the high sensitivity of the FPVS-EEG approach, we corroborate several previous findings from Dzhelyova et al., 2017, in particular significant expression-change responses for all emotions despite short recording times (i.e., 80 s of stimulation were collected for an emotion expressed at 100% of intensity) and significant expression detection in every individual brain, here already found from 60% of expression intensity. There is a right hemispheric advantage for the expression-change response in line with the well-known right-sided asymmetry from lesion studies and neuroimaging (Etcoff, 1984; Sato et al., 2004; Tsuchiya et al., 2008). In addition, there is a differential pattern of spatial distribution depending on the emotion expressed. While the response is mainly recorded over occipito-temporal sites for all emotions, it is more focal for fear and sadness and more dorsal for happiness (in line with Dzhelyova et al., 2017).

We also replicate a triphasic signature in time starting from about 100 ms after the change of facial expression and lasting for several hundreds of milliseconds, indicating long integration of facial expression in the brain despite the brief duration of the expression-change. This observation accords with previous studies showing sensitivity to facial expression for several ERP components ranging from around 80 to 500 ms after stimulus-onset (e.g., Batty and Taylor, 2003; Luo et al., 2010b; Schupp et al., 2004; Smith et al., 2013; Williams et al., 2006; Calvo and Nummenmaa, 2015 for review), although it is fair to say that there is a large degree of inconsistency across these studies contrasting absolute electrophysiological responses to various facial expressions (Calvo and Nummenmaa, 2015). Studies using oddball paradigms generating visual mismatch responses have more directly explored the EEG response to a change of expression by introducing rare deviant facial expressions among frequent standard expressions (e.g., Astikainen et al., 2013; Astikainen and Hietanen, 2009; Kimura et al., 2012; Li et al., 2012; Stefanics et al., 2012; Zhao and Li, 2006). However, the mismatch responses to deviant expression roughly correspond to increased posterior responses at variable early/mid latencies with low consistency across studies (Dzhelyova et al., 2017 for a detailed discussion). This is probably due to several sources of noise (i.e., post-hoc subtraction of the absolute ERPs to standard stimuli from those to deviants, low occurrence of deviants (i.e., 10–20%), variable and sometimes “long” (i.e., larger than

500 ms) intervals between two face stimuli) exacerbated by the low SNR of the ERP approach. In contrast here, no subtraction is required since the expression-change responses reflect direct *differential* activities primarily expressed in the frequency-domain at known narrowband frequencies. Moreover, the fast mode of presentation allows many repetitions within short recording times (i.e., 96 expression-changes within 80 s of stimulation), and stimuli are rapidly masked after single glance perception, reducing the contribution of non-perceptual processes beyond 200 ms post-stimulus. Hence, consistent with Dzhelyova and collaborators' previous findings (Dzhelyova et al., 2017), rather than a single negative deflection, we report a complex pattern of neural responses over time, with two positive and a negative polarity responses, dorsal *vs.* ventral scalp topographies respectively for the first and the second deflections, and a large response over many posterior channels for the third component in response to almost all basic emotions tested.

Interestingly, here, the responses are significant until 406 ms post-stimulus, slightly later than in the previous study (i.e., until 310 ms). The three components peak at 140, 240 and 350 ms after stimulus-onset (compared with 110, 170 and about 250 ms). The latencies observed by Dzhelyova and colleagues were obtained with a squarewave stimulation mode at a rate of 12 Hz, and they found the same components delayed by about 20 ms with a 12 Hz sinusoidal stimulation. Sine-wave stimulation implies that the stimulus is not visible at its onset and contrast progressively increases during the stimulation cycle. By comparing sine *vs.* square wave stimulation modes at a 12 Hz base rate, sinusoidal stimulation produces a delay of about a quarter of a stimulation cycle for all peak latencies (Dzhelyova et al., 2017; Retter and Rossion, 2016a). Here, a quarter of a stimulation cycle corresponds to a duration of about 40 ms. Hence, the 'real' peak latencies may be around 100, 200 and 310 ms post-stimulus, with significant activities until about 366 ms.

Finally, we also find differential patterns of magnitude between emotions with greater amplitudes for disgust, as previously observed (Dzhelyova et al., 2017). Moreover, here, sadness elicits a lower response than any other emotions. In the time-domain, these differences appear from the third component beyond 300 ms post-stimulus. Importantly, these differences in amplitude, and more generally the ability to record a significant response to any brief change of expression, cannot be (fully) accounted by the detection of physical changes between expressive and neutral faces by low-level visual areas. First, random changes of stimulus size occurred at every stimulation cycle, so that local physical differences detectable by visual regions with small receptive fields do not only occur during a change of expression but at every picture onset, projecting activities at the base rate. Second, Dzhelyova et al. (2017) showed that the response is drastically reduced by picture-plane inversion. Third, physical differences between neutral and expressive face pictures estimated by a dissimilarity index do not fully explain the *Emotion* effect. Fourth, as more specifically discussed in the next section, while we used a linear increase in terms of physical differences between morph levels, the increase of expression intensity along the stimulation sequence elicits a complex mixture of linear and non-linear modulations of amplitude. Altogether, these observations largely support an important contribution of high-level perceptual processes in the recorded response to brief changes of facial expression.

Tuning functions for automatic detection of facial expression in the human brain

The capacity of the visual system to detect subtle changes of facial expression and increasingly respond to increasing expression intensity is evidenced here, from a nearly absence of signal when emotions are expressed at 20% of intensity to a highly significant expression-change response at 100% of intensity. Importantly, this effect is clearly dissociated from the effect observed for the base response, which is larger when emotions are expressed at 20% of intensity and decreases when intensity increases, mostly between 20% and 40% of expression intensity. Since the intensity increase is designed along a sweep stimulation sequence of 100 s, the effect of expression intensity also includes the effect of the stimulation period (i.e., from 20% of intensity = the first 20 s of the sequence to 100% of intensity = the last 20 s of the sequence). In other words, the base response is larger at the start of the stimulation sequence and decreases along the sequence, mainly during the first 20 s, while the expression-change response is absent at the start of the sequence and progressively increases along the sequence. This dissociation implies that the two responses largely reflect distinct neural processes. The base response is a general visual response reflecting the brain synchronization to the rapid stream of stimulation, capturing low- and higher-level visual processes elicited at 6 Hz and common to neutral and expressive faces. Its decrease may be explained by rapid adaptation to visual cues repeating at 6 Hz (e.g., individual face identity; Nemrodov et al., 2015; Retter and Rossion, 2016b) and/or by reduced level of attention to the stimulation (Kim et al., 2007; Morgan et al., 1996). In contrast, the expression-change response provides a direct differential measure of the automatic detection of a brief change of facial expression from a neutral face that progressively increases as a function of expression intensity.

Strikingly, the effect of expression intensity on the expression-change response is observed for all emotions and encompasses all components in the time-domain, showing that all visual processes involved in the coding of various facial expressions are sensitive to expression intensity. In contrast, previous ERP research investigating facial expression perception as a function of intensity provided inconsistent observations about the advent of intensity effects throughout the processing stream or the generalizability *vs.* specificity across emotion categories. While some studies suggest early (i.e., P1, around 80 ms) sensitivity to expression intensity (Utama et al., 2009; Wang et al., 2013), others only found later (i.e., N170, around 160 ms) effects (Leppänen et al., 2007; Sprengelmeyer and Jentzsch, 2006). Even later (i.e., EPN, until 600 ms) modulations by expression intensity were described (Sprengelmeyer and Jentzsch, 2006), while no study observed an intensity effect for three components or more in the time-course of visual processing. In addition, some studies observed an effect of intensity for both positive and negative emotions (Utama et al., 2009) whereas others only found it for a negative emotion (Leppänen et al., 2007). Hence, the high sensitivity of the FPVS-EEG approach may be key in revealing an effect of expression intensity for all emotions at the level of three clear components ranging from 80 to 400 ms post-stimulus.

Interestingly, we observed dissociations between different facial emotions. Notably, the detection of brief changes of facial expression from neutrality is obtained for most of the emotions at 40% of intensity, but at 60% of intensity for sadness. Sadness is also the only emotion for which the intensity effect is not significant beyond 300 ms post-stimulus (i.e., for the third component in the time-domain). These specific observations may be explained by less prototypical or low-expressivity full-blown expressions of sadness in our stimulus set that may have reduced its discriminability from neutrality. Indeed, the physical dissimilarity index calculated on our stimuli showed a lower dissimilarity for sadness. However, it was close to the index obtained for happiness, which is associated with a large intensity effect. Another explanation may be that sadness is typically expressed with less salient changes in facial features, as suggested in a previous psychophysical study that also found higher detection threshold for this emotion (Marneweck et al., 2013). Future studies should further explore and clarify this specific pattern found for sadness.

Critically, our approach objectively quantifies automatic detection of brief changes of facial expression depending on intensity directly in the brain with no related behavioral task. This is a considerable advantage because the brain response is free from decisional biases, ceiling effects are reduced (i.e., the response increases for every expression intensity steps) while frequently observed with behavioral accuracy (Utama et al., 2009) or response times (Leppänen et al., 2007), and the estimation of the tuning functions is unconstrained by a specific task (i.e., emotion categorization *vs.* intensity rating revealing abrupt response *vs.* linear increment, see Calder et al., 2000a; Etcoff and Magee, 1992; Hess et al.,

1997; Leleu et al., 2016; Utama et al., 2009). In addition, detection thresholds and tuning functions are not estimated from a mixture of perceptual and post-perceptual processes. This provides a valuable tool to, for instance, determine whether impaired processing of facial expression in clinical populations originates from perceptual *vs.* executive processing (e.g., Leleu et al., 2016). That is, all populations can be tested with the same parameters, even difficult-to-test participants whose cognitive profile strongly hampers their ability to produce adequate explicit behavior. Finally, the ability to determine detection thresholds in a single brain is a clear benefit. Here, we show that already 28% of participants present a significant expression-change response at 20% of intensity, another 39% detect expression-changes at 40% of intensity, and all at 60%, 80% and 100% of intensity. Thus, FPVS-EEG could be used for individual assessment, opening an avenue to systematically evaluate individual differences in perceptual abilities or even to diagnose visual processing disorders.

For the first time, we characterize directly at the brain level the shape of the tuning functions to facial expression of increasing intensity as a combination of gradual sensitivity to expression-change and abrupt emotion categorization. This pattern is subtended over time by two components increasing monotonically with increasing discriminability of expression, the non-linear response profile additionally emerging beyond 300 ms post-stimulus for a third component. From these temporal dynamics, we can initiate a clarification of the functional properties of each component in a three-phase model of the visual processing of facial expression. For the two first components, their similar linear profiles in response to intensity suggest two processes that code facial expression as a function of physical dissimilarity from neutrality, upstream to any categorization process. The first component, which has an early onset and is more dorsal and posterior than the second component, may predominantly involve low-level visual areas that process local physical changes. In addition, it is significant at a greater intensity level than the second component. One explanation may be that low-level visual regions with small receptive fields also respond to physical changes from size variation occurring at each stimulation cycle (i.e., base frequency at 6 Hz), masking those from a low-intensity change of expression. Increased expression intensity may thus be needed to stimulate these areas with an expression-change despite size changing. In line with this account, a previous study explored individual face discrimination using similar periodic stimulation and manipulated size variation at the base rate in a parametric fashion (Dzhelyova and Rossion, 2014). Increased amplitude with increased size variation was found for the base response whereas reduced amplitude, and finally absence of signal, was observed for the first component of the identity-change response. This clearly reveals the contribution of low-level processes within the early time-range for discriminating face identity that can be easily transposed to the time-course of expression-change detection. We also note that the first component seems less sensitive to picture-plane inversion than the following components in Dzhelyova and collaborators' study (2017; Fig. 5), low-level visual information being preserved between upright and upside-down faces.

Perceptual coding may therefore operate at the level of the second component, more laterally and ventrally distributed than the first component, suggesting the contribution of higher-level visual areas. It may reflect the processing of facial movements within the superior temporal sulcus (Puce et al., 2003; Srinivasan et al., 2016) known to increase the amplitude of electrocortical responses around 200 ms post-stimulus (Puce et al., 2003; Rossi et al., 2014). It is also the first component that seems largely reduced by inversion in Dzhelyova et al. (2017), suggesting that the processing of a change induced by facial expression in the whole face configuration, altered for inverted faces (Calder et al., 2000b), occurs within this time-range. Although future research should more precisely disentangle the functional aspects of this component, these findings suggest that the expression-change specific response occurring about 200 ms post-stimulus signs the first high-level perceptual processing of facial expressions.

Therefore, downstream to this first perceptual stage, the third component is in an ideal position to reflect more elaborate processing facilitating the categorization of the emotional content of facial expression. Contrary to the two first components, its amplitude increase as a function of intensity is a mixture of linear, quadratic and cubic functions indicating more abrupt modulation by a change of expression. It is only significant up to 60% of expression intensity and about 75% of its amplitude increase arises beyond this threshold. Together with the fact that magnitude differences between emotions are only present for this component, this tuning pattern strongly suggests a categorical process responding in an all-or-none fashion when an emotion category is attributed to facial expression. The underlying neural mechanisms may be, for instance, the integration of affective value within the visual processing of facial expression, potentially through reentrant connections between the visual system and heteromodal "emotional" regions such as the amygdala (Vuilleumier et al., 2004). This explanation is consistent with a late (280–410 ms) activation within the amygdala (Luo et al., 2010a) and with the absence of a late (500–600 ms) ERP effect of expressive *vs.* neutral faces in patients with amygdala damage (Rotshtein et al., 2010). Interestingly, using morphs within one or between two emotion categories with similar amount of physical differences in both conditions, a recent fMRI study revealed that the STS and the amygdala respectively respond to facial expression in a continuous (linear) *vs.* categorical (abrupt) fashion (Harris et al., 2012). This finding thus supports our interpretation of linear perceptual coding of expression within the STS indexed by the second component around 200 ms post-stimulus, followed by emotion categorization from integrated visual-affective processing through feedback from the amygdala to the cortex beyond 300 ms.

Limitations

A first limitation of the present study comes from the use of morph continua to assess the perception of facial expression as a function of intensity. That is, since they assume equivalent contribution and speed of motion of all facial action units (i.e., facial movements composing a facial expression), linear morphs between neutral and full-blown expressive faces are not the most ecological way to simulate subtle facial expressions. Indeed, in natural situations, some facial actions may start to be involved at low intensities while others may rather be active only from a high level of expression. For instance, a subtle happy face may only imply that lip corners pull up (i.e., smiling lips) while the cheeks rise and the muscles around the eyes contract only for more intense happy expressions. In addition, simulation of variable intensity from linear morphs may be more or less artificial across emotion categories and some differences between emotions may be confounded with a more or less ecological simulation of expression intensity. We acknowledge that this limitation should be addressed in future studies by using, for instance, natural pictures of individual faces expressing an emotion at each intensity step tested. Nevertheless, as in many studies that previously investigated this issue (e.g., Calder et al., 2000a; Etcoff and Magee, 1992; Gao and Maurer, 2010; Hess et al., 1997; Leleu et al., 2016; Marneweck et al., 2013), we note that one clear advantage of using linear morphs to explore facial expression perception as a function of intensity is a well-controlled parametric manipulation of the physical changes across intensity steps. This choice was motivated in the present study by our purpose of identifying non-linear brain activities in response to linear physical changes as a signature of categorical perception of facial expression. Hence, we consider our present observations as an essential prerequisite for future investigations aiming at delineating facial expression perception at various intensities using more ecological stimuli.

Another limitation comes from the use of an oddball-like paradigm in which the repetition of a standard/frequent neutral face is interspersed by a deviant/rare facial expression. As in typical oddball ERP paradigms targeting visual mismatch responses (Kimura, 2012 for review), in

particular for facial expressions (e.g., Astikainen et al., 2013; Astikainen and Hietanen, 2009; Kimura et al., 2012; Li et al., 2012; Stefanics et al., 2012; Zhao and Li, 2006), the expression-change response may be contaminated by, or even mainly reflect, a general mechanism of repetition suppression or deviancy detection. In other words, one could argue that what is defined here as an "expression-change response" is rather driven by an unspecific neural process of change detection unrelated to the visual coding of facial expression. While we acknowledge that the neural mechanisms subtending the response to facial expression-change remain unknown in our study, and that the various theoretical interpretations of the functional property of visual mismatch responses such as repetition suppression/release from adaptation or regularity violation/deviancy detection (Garrido et al., 2009 for review) may also apply to the present observations, we believe that several elements support an interpretation in terms of specific response to the detection of a change of facial expression. First, the response focuses on the right occipito-temporal cortex, a signature of face-specific high-level visual activity. Second, the pattern of response across facial emotions and morph levels is not solely explained by physical changes between neutral and expressive faces, as indicated by the analyses on data normalized by a physical dissimilarity index and by the complex combination of linear and non-linear brain activities dissociated in time, thus pointing to the contribution of various visual processes including high-level perceptual processes. Third, the discrimination response is largely reduced by picture-plane inversion (Dzhelyova et al., 2017) which cannot be accounted for by an unspecific mechanism detecting deviancies, identical for upright and inverted faces. These observations are hardly reconcilable with an undifferentiated change detection process, but rather support content-specific high-level perceptual processes. Finally, we note that most interpretations of visual mismatch responses consider the neural coding of sensory information directly related to the content of information, irrespective of the precise mechanism appealed to explain this coding (e.g., Czigler, 2014; Winkler and Czigler, 2012). In sum, while we consider that "release from adaptation" or "deviancy detection" are different theoretical constructs that can provide a general account of the neural responses measured in an FPVS oddball paradigm, and of how sensory information is processed in the brain in general, such processing is intrinsically related to the nature and content of sensory information manipulated in a specific design.

According to a predictive coding framework (e.g., Garrido et al., 2008; Kimura, 2012; Stefanics et al., 2014), one could also argue that the expression-change brain response identified here is generated or contaminated by expectation-related processes since changes of facial expression are strictly periodic (i.e., 1 out of 5 faces) and thus temporally predictable within FPVS sequences. However, it should be noted that participants in the study were unable to notice any periodicity in the expression changes, given the high rate of presentation. Moreover, the standard neutral face stimuli are variable, since there are substantial changes of size at every stimulation cycle. In addition, one would have to account for non-linear increases of prediction signals as a function of facial expression intensity. Finally, and perhaps most convincingly, there is recent evidence that face categorization EEG responses elicited in FPVS paradigms are immune to temporal predictability (Quek and Rossion, 2017). In that study, face-selective responses did not differ in magnitude and spatio-temporal characteristics whether the stimuli appeared periodically or not (i.e., no effect of predictability). In addition, high temporal expectations of encountering a periodically presented face were insufficient to elicit any neural response in rare trials where a face was absent (Quek and Rossion, 2017). While we acknowledge that further studies should generalize these observations of immunity to temporal expectation to other domains such as the detection of brief changes of expression, this does not prevent the ability of the current approach to provide a sensitive and specific measure of facial expression perception that will help future work to draw a clear picture of the visual processing of facial expression.

Conclusions

Using FPVS-EEG and a sweep-VEP design manipulating expression intensity in a parametric fashion, we extend previous findings about the human brain ability to automatically detect changes of facial expression from a neutral face at a glance, and additionally quantify detection thresholds as a function of expression intensity over right occipito-temporal cortex, even at the individual level. For almost all tested emotions, we identify early gradual sensitivity to expression intensity, from about 100 ms after stimulus-onset, followed by a more categorical tuning function beyond 300 ms post-stimulus. These findings provide a testable functional model for the visual coding of facial expression changes and its temporal dynamics that will help future research to fully characterize the visual mechanisms underlying facial expression perception. From a broader extent, they show that FPVS-EEG constitutes a unique approach for objective quantification of facial expression processing in various populations. Direct neural assessment of detection thresholds and tuning functions depending on intensity is thus made possible in difficult-to-test individuals (e.g., with high intellectual disability) and can even be used in a clinical context for prognostic or diagnostic purpose.

Acknowledgments

The authors thank three anonymous Reviewers for their constructive comments on previous versions of the paper. The authors are grateful to the people who participated in the study and to Romain Patrux for help in data collection. This work received support from the "Conseil Régional Bourgogne Franche-Comté" (PARI grant to JYB, AL, KD, RB and BS, and FABER grant to AL), the FEDER (European Funding for Regional Economic Development), the French "Investissements d'Avenir" program, project ISITE-BFC (contract ANR-15-IDEX-03 to JYB), an "Action de Recherche Concertée" grant (ARC; 13/18-053) and the Louvain Foundation for BR and MD. The authors declare no competing interest.

Appendix A. Supplementary data

Supplementary data related to this article can be found at https://doi.org/10.1016/j.neuroimage.2018.06.048.

Preprocessed EEG data associated with this article are freely available at https://doi.org/10.12751/g-node.2ff1af.

References

Adrian, E.D., Matthews, B.H.C., 1934. The Berger rhythm: potential changes from the occipital lobes in man. Brain 57, 355–385. https://doi.org/10.1093/brain/57.4.355.

Ales, J.M., Farzin, F., Rossion, B., Norcia, A.M., 2012. An objective method for measuring face detection thresholds using the sweep steady-state visual evoked response. J. Vis. 12, 18–18. https://doi.org/10.1167/12.10.18.

Astikainen, P., Cong, F., Ristaniemi, T., Hietanen, J., 2013. Event-related potentials to unattended changes in facial expressions: detection of regularity violations or encoding of emotions? Front. Hum. Neurosci. 7. https://doi.org/10.3389/fnhum.2013.00557.

Astikainen, P., Hietanen, J.K., 2009. Event-related potentials to task-irrelevant changes in facial expressions. Behav. Brain Funct. 5 (30). https://doi.org/10.1186/1744-9081-5-30.

Batty, M., Taylor, M.J., 2003. Early processing of the six basic facial emotional expressions. Cognit. Brain Res. 17, 613–620.

Bombari, D., Schmid, P.C., Schmid Mast, M., Birri, S., Mast, F.W., Lobmaier, J.S., 2013. Emotion recognition: the role of featural and configural face information. Q. J. Exp. Psychol. 66, 2426–2442. https://doi.org/10.1080/17470218.2013.789065.

Calder, A.J., Rowland, D., Young, A.W., Nimmo-Smith, I., Keane, J., Perrett, D.I., 2000a. Caricaturing facial expressions. Cognition 76, 105–146. https://doi.org/10.1016/S0010-0277(00)00074-3.

Calder, A.J., Young, A.W., Keane, J., Dean, M., 2000b. Configural information in facial expression perception. J. Exp. Psychol. Hum. Percept. Perform. 26, 527–551. https://doi.org/10.1037/0096-1523.26.2.527.

Calder, A.J., Young, A.W., Perrett, D.I., Etcoff, N.L., Rowland, D., 1996. Categorical perception of morphed facial expressions. Vis. Cognit. 3, 81–118.

Calvo, M.G., Avero, P., Fernández-Martín, A., Recio, G., 2016. Recognition thresholds for static and dynamic emotional faces. Emotion 16, 1186–1200. https://doi.org/10.1037/emo0000192.

Calvo, M.G., Nummenmaa, L., 2015. Perceptual and affective mechanisms in facial expression recognition: an integrative review. Cognit. Emot. 30, 1081–1106. https://doi.org/10.1080/02699931.2015.1049124.
Carroll, J.M., Russell, J.A., 1997. Facial expressions in Hollywood's protrayal of emotion. J. Pers. Soc. Psychol. 72, 164.
Chambon, V., Baudouin, J.-Y., Franck, N., 2006. The role of configural information in facial emotion recognition in schizophrenia. Neuropsychologia 44, 2437–2444.
Czigler, I., 2014. Visual mismatch negativity and categorization. Brain Topogr. 27, 590–598. https://doi.org/10.1007/s10548-013-0316-8.
Darwin, C., 1872. The Expression of the Emotions in Man and Animals, first. ed. John Murray, London.
Durand, K., Gallay, M., Seigneuric, A., Robichon, F., Baudouin, J.-Y., 2007. The development of facial emotion recognition: the role of configural information. J. Exp. Child Psychol. 97, 14–27.
Dzhelyova, M., Jacques, C., Rossion, B., 2017. At a single glance: fast periodic visual stimulation uncovers the spatio-temporal dynamics of brief facial expression changes in the human brain. Cereb. Cortex 27, 4106–4123. https://doi.org/10.1093/cercor/bhw223.
Dzhelyova, M., Rossion, B., 2014. The effect of parametric stimulus size variation on individual face discrimination indexed by fast periodic visual stimulation. BMC Neurosci. 15, 87.
Ekman, P., 1992. An argument for basic emotions. Cognit. Emot. 6, 169–200.
Elfenbein, H.A., Ambady, N., 2002. On the universality and cultural specificity of emotion recognition: a meta-analysis. Psychol. Bull. 128, 203–235. https://doi.org/10.1037//0033-2909.128.2.203.
Etcoff, N.L., 1984. Perceptual and conceptual organization of facial emotions: hemispheric differences. Brain Cognit. 3, 385–412. https://doi.org/10.1016/0278-2626(84)90030-7.
Etcoff, N.L., Magee, J.J., 1992. Categorical perception of facial expressions. Cognition 44, 227–240. https://doi.org/10.1016/0010-0277(92)90002-Y.
Gao, X., Maurer, D., 2010. A happy story: developmental changes in children's sensitivity to facial expressions of varying intensities. J. Exp. Child Psychol. 107, 67–86. https://doi.org/10.1016/j.jecp.2010.05.003.
Garrido, M.I., Friston, K.J., Kiebel, S.J., Stephan, K.E., Baldeweg, T., Kilner, J.M., 2008. The functional anatomy of the MMN: a DCM study of the roving paradigm. Neuroimage 42, 936–944. https://doi.org/10.1016/j.neuroimage.2008.05.018.
Garrido, M.I., Kilner, J.M., Stephan, K.E., Friston, K.J., 2009. The mismatch negativity: a review of underlying mechanisms. Clin. Neurophysiol. 120, 453–463. https://doi.org/10.1016/j.clinph.2008.11.029.
Harris, R.J., Young, A.W., Andrews, T.J., 2012. Morphing between expressions dissociates continuous from categorical representations of facial expression in the human brain. Proc. Natl. Acad. Sci. Unit. States Am. 109, 21164–21169. https://doi.org/10.1073/pnas.1212207110.
Hess, U., Blairy, S., Kleck, R.E., 1997. The intensity of emotional facial expressions and decoding accuracy. J. Nonverbal Behav. 21, 241–257.
Horstmann, G., 2002. Facial expressions of emotion: does the prototype represent central tendency, frequency of instantiation, or an ideal? Emotion 2, 297–305.
Jack, R.E., Garrod, O.G.B., Yu, H., Caldara, R., Schyns, P.G., 2012. Facial expressions of emotion are not culturally universal. Proc. Natl. Acad. Sci. Unit. States Am. 109, 7241–7244. https://doi.org/10.1073/pnas.1200155109.
Kim, H., Somerville, L., Johnstone, T., Polis, S., Alexander, A., Shin, L., Whalen, P., 2004. Contextual modulation of amygdala responsivity to surprised faces. J. Cognit. Neurosci. 16, 1730–1745.
Kim, Y.J., Grabowecky, M., Paller, K.A., Muthu, K., Suzuki, S., 2007. Attention induces synchronization-based response gain in steady-state visual evoked potentials. Nat. Neurosci. 10, 117–125. https://doi.org/10.1038/nn1821.
Kimura, M., 2012. Visual mismatch negativity and unintentional temporal-context-based prediction in vision. Int. J. Psychophysiol. 83, 144–155. https://doi.org/10.1016/j.ijpsycho.2011.11.010.
Kimura, M., Kondo, H., Ohira, H., Schröger, E., 2012. Unintentional temporal context–based prediction of emotional faces: an electrophysiological study. Cereb. Cortex 22, 1774–1785. https://doi.org/10.1093/cercor/bhr244.
Leleu, A., Demily, C., Franck, N., Durand, K., Schaal, B., Baudouin, J.-Y., 2015. The odor context facilitates the perception of low-intensity facial expressions of emotion. PLoS One 10, e0138656. https://doi.org/10.1371/journal.pone.0138656.
Leleu, A., Saucourt, G., Rigard, C., Chesnoy, G., Baudouin, J.-Y., Rossi, M., Edery, P., Franck, N., Demily, C., 2016. Facial emotion perception by intensity in children and adolescents with 22q11.2 deletion syndrome. Eur. Child Adolesc. Psychiatr. 25, 297–310. https://doi.org/10.1007/s00787-015-0741-1.
Leppänen, J.M., Kauppinen, P., Peltola, M.J., Hietanen, J.K., 2007. Differential electrocortical responses to increasing intensities of fearful and happy emotional expressions. Brain Res. 1166, 103–109. https://doi.org/10.1016/j.brainres.2007.06.060.
Li, X., Lu, Y., Sun, G., Gao, L., Zhao, L., 2012. Visual mismatch negativity elicited by facial expressions: new evidence from the equiprobable paradigm. Behav. Brain Funct. 8 (7). https://doi.org/10.1186/1744-9081-8-7.
Liu-Shuang, J., Norcia, A.M., Rossion, B., 2014. An objective index of individual face discrimination in the right occipito-temporal cortex by means of fast periodic oddball stimulation. Neuropsychologia 52, 57–72. https://doi.org/10.1016/j.neuropsychologia.2013.10.022.
Luck, S.J., 2005. An Introduction to the Event-related Potential Technique. MIT Press, Cambridge, MA.
Luo, Q., Holroyd, T., Majestic, C., Cheng, X., Schechter, J., Blair, R.J., 2010a. Emotional automaticity is a matter of timing. J. Neurosci. 30, 5825–5829. https://doi.org/10.1523/JNEUROSCI.BC-5668-09.2010.
Luo, W., Feng, W., He, W., Wang, N.-Y., Luo, Y.-J., 2010b. Three stages of facial expression processing: ERP study with rapid serial visual presentation. Neuroimage 49, 1857–1867. https://doi.org/10.1016/j.neuroimage.2009.09.018.
Makeig, S., Bell, A.J., Jung, T.-P., Sejnowski, T.J., 1996. Independent component analysis of electroencephalographic data. In: Touretzky, D., Mozer, M., Hasselmo, M. (Eds.), Advances in Neural Information Processing Systems. MIT Press, Cambridge, MA, pp. 145–151.
Marneweck, M., Loftus, A., Hammond, G., 2013. Psychophysical measures of sensitivity to facial expression of emotion. Front. Psychol. 4. https://doi.org/10.3389/fpsyg.2013.00063.
McCarthy, G., Wood, C.C., 1985. Scalp distributions of event-related potentials: an ambiguity associated with analysis of variance models. Electroencephalogr. Clin. Neurophysiol. Potentials Sect 62, 203–208. https://doi.org/10.1016/0168-5597(85)90015-2.
Miellet, S., Vizioli, L., He, L., Zhou, X., Caldara, R., 2013. Mapping face recognition information use across cultures. Front. Psychol. 4. https://doi.org/10.3389/fpsyg.2013.00034.
Morgan, S.T., Hansen, J.C., Hillyard, S.A., 1996. Selective attention to stimulus location modulates the steady-state visual evoked potential. Proc. Natl. Acad. Sci. Unit. States Am. 93, 4770–4774.
Motley, M.T., Camden, C.T., 1988. Facial expression of emotion: a comparison of posed expressions versus spontaneous expressions in an interpersonal communication setting. West. J. Commun. Incl. Commun. Rep 52, 1–22.
Näätänen, R., Gaillard, A.W.K., Mäntysalo, S., 1978. Early selective-attention effect on evoked potential reinterpreted. Acta Psychol. 42, 313–329. https://doi.org/10.1016/0001-6918(78)90006-9.
Nemrodov, D., Jacques, C., Rossion, B., 2015. Temporal dynamics of repetition suppression to individual faces presented at a fast periodic rate. Int. J. Psychophysiol. 98, 35–43. https://doi.org/10.1016/j.ijpsycho.2015.06.006.
Norcia, A.M., Appelbaum, L.G., Ales, J.M., Cottereau, B.R., Rossion, B., 2015. The steady-state visual evoked potential in vision research: a review. J. Vis. 15, 4–4. https://doi.org/10.1167/15.6.4.
Puce, A., Syngeniotis, A., Thompson, J.C., Abbott, D.F., Wheaton, K.J., Castiello, U., 2003. The human temporal lobe integrates facial form and motion: evidence from fMRI and ERP studies. Neuroimage 19, 861–869. https://doi.org/10.1016/S1053-8119(03)00189-7.
Quek, G.L., Rossion, B., 2017. Category-selective human brain processes elicited in fast periodic visual stimulation streams are immune to temporal predictability. Neuropsychologia 104, 182–200. https://doi.org/10.1016/j.neuropsychologia.2017.08.010.
Regan, D., 1989. Human Brain Electrophysiology: Evoked Potentials and Evoked Magnetic fields in Science and Medicine. Elsevier, New York.
Regan, D., 1973. Rapid objective refraction using evoked brain potentials. Invest. Ophthalmol. Vis. Sci. 12, 669–679.
Retter, T.L., Rossion, B., 2016a. Uncovering the neural magnitude and spatio-temporal dynamics of natural image categorization in a fast visual stream. Neuropsychologia 91, 9–28. https://doi.org/10.1016/j.neuropsychologia.2016.07.028.
Retter, T.L., Rossion, B., 2016b. Visual adaptation provides objective electrophysiological evidence of facial identity discrimination. Cortex 80, 35–50. https://doi.org/10.1016/j.cortex.2015.11.025.
Rossi, A., Parada, F.J., Kolchinsky, A., Puce, A., 2014. Neural correlates of apparent motion perception of impoverished facial stimuli: a comparison of ERP and ERSP activity. Neuroimage 98, 442–459. https://doi.org/10.1016/j.neuroimage.2014.04.029.
Rossion, B., Torfs, K., Jacques, C., Liu-Shuang, J., 2015. Fast periodic presentation of natural images reveals a robust face-selective electrophysiological response in the human brain. J. Vis. 15 (18). https://doi.org/10.1167/15.1.18.
Rotshtein, P., Richardson, M.P., Winston, J.S., Kiebel, S.J., Vuilleumier, P., Eimer, M., Driver, J., Dolan, R.J., 2010. Amygdala damage affects event-related potentials for fearful faces at specific time windows. Hum. Brain Mapp. 31, 1089–1105. https://doi.org/10.1002/hbm.20921.
Sato, W., Kochiyama, T., Yoshikawa, S., Naito, E., Matsumura, M., 2004. Enhanced neural activity in response to dynamic facial expressions of emotion: an fMRI study. Cognit. Brain Res. 20, 81–91.
Sauter, D.A., Eisner, F., 2013. Commonalities outweigh differences in the communication of emotions across human cultures. Proc. Natl. Acad. Sci. Unit. States Am. 110. E180–E180. https://doi.org/10.1073/pnas.1209522110.
Schupp, H.T., Öhman, A., Junghöfer, M., Weike, A.I., Stockburger, J., Hamm, A.O., 2004. The facilitated processing of threatening faces: an ERP analysis. Emotion 4, 189–200. https://doi.org/10.1037/1528-3542.4.2.189.
Smith, E., Weinberg, A., Moran, T., Hajcak, G., 2013. Electrocortical responses to NIMSTIM facial expressions of emotion. Int. J. Psychophysiol. 88, 17–25. https://doi.org/10.1016/j.ijpsycho.2012.12.004.
Sprengelmeyer, R., Jentzsch, I., 2006. Event related potentials and the perception of intensity in facial expressions. Neuropsychologia 44, 2899–2906. https://doi.org/10.1016/j.neuropsychologia.2006.06.020.
Srinivasan, R., Golomb, J.D., Martinez, A.M., 2016. A neural basis of facial action recognition in humans. J. Neurosci. 36, 4434–4442. https://doi.org/10.1523/JNEUROSCI.1704-15.2016.
Stefanics, G., Csukly, G., Komlósi, S., Czobor, P., Czigler, I., 2012. Processing of unattended facial emotions: a visual mismatch negativity study. Neuroimage 59, 3042–3049. https://doi.org/10.1016/j.neuroimage.2011.10.041.
Stefanics, G., Kremláček, J., Czigler, I., 2014. Visual mismatch negativity: a predictive coding view. Front. Hum. Neurosci. 8 (666). https://doi.org/10.3389/fnhum.2014.00666.

Tsuchiya, N., Kawasaki, H., Oya, H., Iii, M.A.H., Adolphs, R., 2008. Decoding face information in time, frequency and space from direct intracranial recordings of the human brain. PLoS One 3, e3892. https://doi.org/10.1371/journal.pone.0003892.

Utama, N.P., Takemoto, A., Koike, Y., Nakamura, K., 2009. Phased processing of facial emotion: an ERP study. Neurosci. Res. 64, 30–40. https://doi.org/10.1016/j.neures.2009.01.009.

Vernet, M., Baudouin, J.-Y., Franck, N., 2008. Facial emotion space in schizophrenia. Cognit. Neuropsychiatry 13, 59–73. https://doi.org/10.1080/13546800701795228.

Vuilleumier, P., Richardson, M.P., Armony, J.L., Driver, J., Dolan, R.J., 2004. Distant influences of amygdala lesion on visual cortical activation during emotional face processing. Nat. Neurosci. 7, 1271–1278.

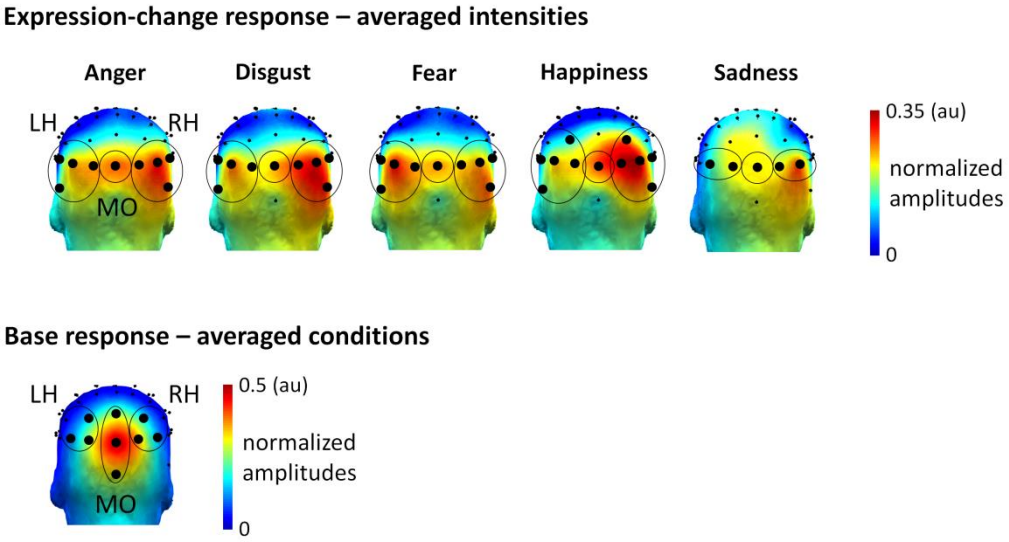
Wang, Y., Liu, F., Li, R., Yang, Y., Liu, T., Chen, H., 2013. Two-stage processing in automatic detection of emotional intensity: a scalp event-related potential study. Neuroreport 24, 818–821. https://doi.org/10.1097/WNR.0b013e328364d59d.

Williams, L.M., Palmer, D., Liddell, B.J., Song, L., Gordon, E., 2006. The 'when' and 'where' of perceiving signals of threat versus non-threat. Neuroimage 31, 458–467. https://doi.org/10.1016/j.neuroimage.2005.12.009.

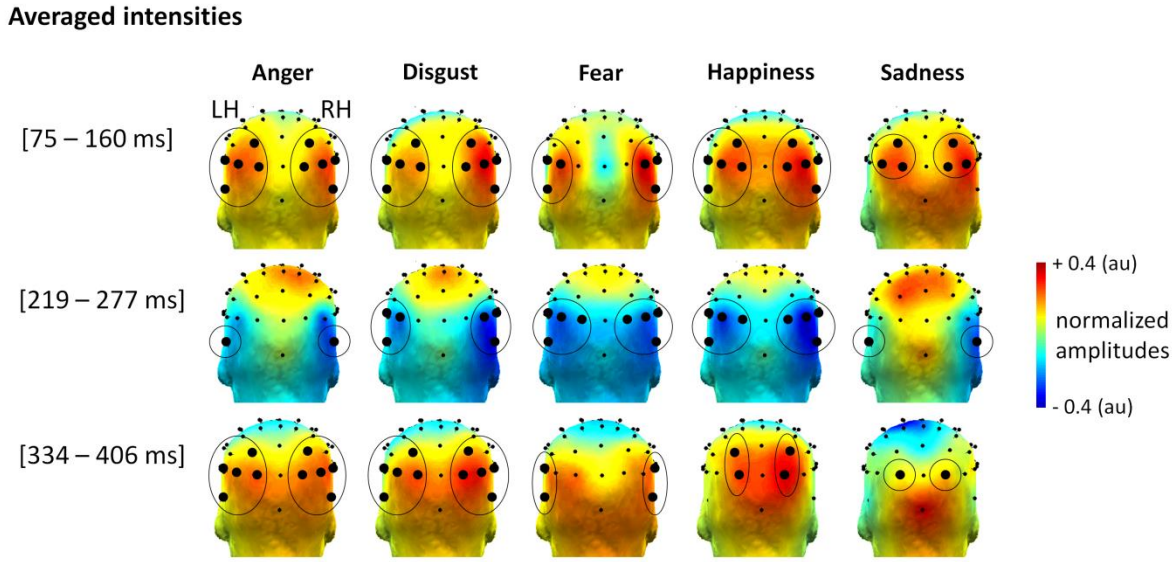
Winkler, I., Czigler, I., 2012. Evidence from auditory and visual event-related potential (ERP) studies of deviance detection (MMN and vMMN) linking predictive coding theories and perceptual object representations. Int. J. Psychophysiol. 83, 132–143. https://doi.org/10.1016/j.ijpsycho.2011.10.001.

Young, A.W., Rowland, D., Calder, A.J., Etcoff, N.L., Seth, A., Perrett, D.I., 1997. Facial expression megamix: tests of dimensional and category accounts of emotion recognition. Cognition 63, 271–313.

Zhao, L., Li, J., 2006. Visual mismatch negativity elicited by facial expressions under non-attentional condition. Neurosci. Lett. 410, 126–131. https://doi.org/10.1016/j.neulet.2006.09.081.



Supplementary Fig. S1. Frequency-domain: definition of the regions-of-interest (ROIs) for statistical analysis. 3-D topographical maps (posterior view) of normalized summed baseline-corrected amplitudes (BCA) for the expression-change response averaged across intensities for each emotion (top) and for the base response averaged across all conditions (bottom). A data-driven approach (see section 2.5.2.1.) was applied to determine three regions-of-interest (ROIs) to include in statistical analysis for each brain response. For both responses, ROIs are highlighted in topographical maps. *For the expression-change response*, ROIs were determined separately for each emotion pooling data across participants and intensities since different topographies were previously observed depending on the emotion expressed (Dzhelyova et al., 2017). Significant activities were delineated over right occipito-temporal electrodes for anger and disgust (P10, P8, PO8, O2), only three of them being significant for fear (P10, PO8, O2) and only PO8 for sadness. Happiness elicited more dorsal responses with significant responses over PO4 in addition to all previous channels. In the left hemisphere, only O1 was significant for sadness while PO7 was also significant for anger and happiness. PO7 and PO9 were significant for disgust and significant channels encompassed a larger region for fear including P7, P9, PO7 and O1. The medial occipital channel Oz was also significant for all emotions. Based on these results, three ROIs were defined according to the electrodes outlined in the right hemisphere (pooled in the RH ROI), the symmetrical channels in the left hemisphere (LH ROI), and Oz (MO ROI). *For the base response*, ROIs were determined pooling data across participants and conditions. Significant activities were found over medial occipital electrodes (Oz, Iz, POz), and right (O2, PO4, PO8) and left (O1, PO3) occipito-parietal and occipito-temporal electrodes. These channels were respectively averaged in the MO, RH and LH ROIs, the latter additionally considering PO7 to match the channels pooled in RH.



Supplementary Fig. S2. Time-domain: definition of the regions-of-interest (ROIs) for statistical analysis. 3-D topographical maps (posterior view) of mean normalized amplitude (au: arbitrary unit) for each emotion averaged across intensities and for each expression-change specific components (time-windows: 75 – 160 ms, 219 – 277 ms, 334 – 406 ms). As for the frequency-domain analysis, a data-driven approach was applied to determine two regions-of-interest (ROIs) to include in statistical analyses for each component (highlighted in topographical maps). Only lateral electrodes showing a significant expression-change response in the frequency-domain analysis (i.e., O1/2, PO3/4, PO7/8, P7/8, P9/10) were considered. Mean amplitudes were calculated within each time-window and then normalized (McCarthy and Wood, 1985) for each emotion averaged across intensities. *T*-tests against zero were calculated on these normalized mean amplitudes and significant electrodes were averaged in LH and RH ROIs for each emotion and component. *For the first component*, a significant response was identified over all channels for anger, disgust and happiness (P9/10, P7/8, PO7/8, O1/2, PO3/4), only three of them being significant for fear (P9/10, P7/8, PO7/8) and sadness (PO7/8, O1/2, PO3/4). *For the second component*, significant activities were delineated over four electrodes for fear and happiness (P9/10, P7/8, PO7/8, O1/2) while only three electrodes were significant for disgust (P9/10, P7/8, PO7/8) and only one electrode for anger and sadness (P9/10). *For the third component*, a significant response was observed over all channels for anger and disgust (P9/10, P7/8, PO7/8, O1/2, PO3/4), only two of them being significant for fear (P9/10, P7/8) and happiness (O1/2, PO3/4), and only one (O1/2) for sadness.