

# Reliability of individual differences in neural face identity discrimination

Lisa Stacchi<sup>a,\*</sup>, Joan Liu-Shuang<sup>b</sup>, Meike Ramon<sup>a,1</sup>, Roberto Caldara<sup>a,1</sup>

<sup>a</sup> Eye and Brain Mapping Laboratory (iBMLab), Department of Psychology, University of Fribourg, Switzerland

<sup>b</sup> Psychological Science Research Institute & Institute of Neuroscience, University of Louvain, Belgium



## ARTICLE INFO

### Keywords:

Individual differences  
Face discrimination  
Reliability  
Viewing positions  
FPVS-EEG

## ABSTRACT

Over the past years, much interest has been devoted to understanding how individuals differ in their ability to process facial identity. Fast periodic visual stimulation (FPVS) is a promising technique to obtain objective and highly sensitive neural correlates of face processing across various populations, from infants to neuropsychological patients. Here, we use FPVS to investigate how neural face identity discrimination varies in amplitude and topography across observers. To ascertain more detailed inter-individual differences, we parametrically manipulated the visual input fixated by observers across ten viewing positions (VPs). Specifically, we determined the *inter-session reliability* of VP-dependent neural face discrimination responses, both across and *within* observers (6-month inter-session interval). All observers exhibited idiosyncratic VP-dependent neural response patterns, with reliable individual differences in terms of response amplitude for the majority of VPs. Importantly, the topographical reliability varied across VPs and observers, the majority of which exhibited reliable responses only for specific VPs. Crucially, this topographical reliability was positively correlated with the response magnitude over occipito-temporal regions: observers with stronger responses also displayed more reliable response topographies. Our data extend previous findings of idiosyncrasies in visuo-perceptual processing. They highlight the need to consider intra-individual neural response reliability in order to better understand the functional role(s) and underlying basis of such inter-individual differences.

## 1. Introduction

Faces convey an abundance of information and our ability to process them efficiently is crucial for social interactions. This continues to motivate a large and growing area of research dedicated to unraveling the principles governing human face cognition. The majority of studies have sought to identify *commonalities* across neurotypical observers. Consequently, individual differences have long been considered as noise, the influence of which should be minimized by averaging. Studies employing such group statistics across a range of methodologies have identified the eye region as a highly diagnostic source of information for facial identity processing: it receives the most fixations (Yarbus, 1967; Henderson et al., 2005) in Western observers (e.g., Blais et al., 2008; for a review see Caldara, 2017), is behaviorally relevant (Schyns et al., 2002), and elicits stronger neural face-sensitive responses (Nemrodov et al., 2014; de Lissa et al., 2014).

However, an increasing number of studies indicate that group-average results are not necessarily representative of the population (Arizpe et al., 2017). Adding to inter-individual variability on the

behavioral level (e.g., Wilhelm et al., 2010; Herzmann et al., 2010; Kaltwasser et al., 2014), recent findings emphasize the importance of idiosyncratic visual sampling strategies. Specifically, fixation patterns exhibited by individual observers suggest reliable preferences in sampling different facial information (Mehouadar et al., 2014; Kanan et al., 2015; Peterson and Eckstein, 2013; see also Walker-Smith et al., 1977).

Efforts have also been made to characterize the neural correlates of individual differences in perceptual abilities. For instance, these studies have shown that subjects' electrophysiological face-sensitive event-related potentials (ERPs) differ in terms of both amplitude and latency (Gaspar et al., 2011; Herzmann et al., 2010; Kaltwasser et al., 2014; Das et al., 2010; see also Turano et al., 2016 for sub-group analyses). Inter-individual variation has also been reported for neural face discrimination responses obtained using a different approach: fast periodic visual stimulation (FPVS). FPVS relies on the brain's tendency to synchronize to the frequency of external stimulation. This is exploited to derive a neural measure of a process of interest. The process of face discrimination has been successfully measured with FPVS in an oddball paradigm, where a sequence of same-identity base stimuli are presented

\* Corresponding author. Department of Psychology, University of Fribourg, Fautigny 2, 1700, Fribourg, Switzerland.

E-mail address: [lisa.stacchi@unifr.ch](mailto:lisa.stacchi@unifr.ch) (L. Stacchi).

<sup>1</sup> Both authors contributed equally to this work.

<https://doi.org/10.1016/j.neuroimage.2019.01.023>

Received 7 August 2018; Received in revised form 19 December 2018; Accepted 9 January 2019

Available online 14 January 2019

1053-8119/© 2019 Elsevier Inc. All rights reserved.

at a constant high frequency rate, with periodically intervening oddball stimuli conveying different identities (e.g., A-A-A-A-A-B-A-A-A-A-A-C...). Neural synchronization to the oddball frequency (the frequency of identity change) provides an implicit, objective and highly sensitive measure of neural face discrimination that has been used in healthy and neuropsychological cohorts (Liu-Shuang et al., 2016; Liu-Shuang et al., 2014; Norcia et al., 2015).

Using this paradigm, Xu et al. (Xu et al., 2017) were the first to report pronounced individual differences in terms of both the amplitude and topography of this neural face discrimination response. Beyond this, the authors reported small but significant correlations with behavioral performance, suggesting that the magnitude of this neural response is at least partially driven by face identity processing abilities. Although their findings further emphasize the functional importance of neural idiosyncrasies, two important questions remain unanswered.

The first question concerns the impact of facial information fixated. Normally, in the context of neuroimaging studies, observers are required to fixate a point on or near the center of the face (e.g., Liu-Shuang et al., 2014; Xu et al., 2017). Standardizing the visual input is a commonplace procedure in pursuit of comparable stimulus-related responses. The aforementioned functional individual differences in visual sampling strategies, however, suggest that the *diagnosticity* of visual information differs across observers. If such observer-dependent preferences for specific facial information or viewing positions (VPs) also existed on the neural level, standardizing the input through enforced fixation on one only VP would be problematic. Specifically, recording the response during fixation of a central VP (e.g., the nose) might accurately represent observers who prefer encoding face-stimuli at this VP, but may misrepresent those who prefer other facial information (e.g., the left eye). To clarify this issue, in our FPVS paradigm we presented whole-face stimuli under enforced fixation to ten different VPs.

Such potential individual differences in VP-preference are directly related to the second open question: To which degree do idiosyncratic FPVS face identity discrimination responses reflect genuine, intrinsic inter-individual differences, as opposed to situational noise? One way to assess this is to determine whether idiosyncratic face discrimination responses reoccur systematically over time. Noise is a random factor, which therefore impacts a measure of interest differently across recordings. Conversely, intrinsic characteristics of a given subject are more stable (Seghier and Price, 2018). As a consequence, their influence on a given response should systematically reoccur over time. Thus, quantifying test-retest reliability across a range of VPs provides a detailed means to estimate the impact of nuisance factors on the signal(s) of interest.

To this date, the reliability of face discrimination responses obtained by FPVS has only been investigated across trials acquired within the same testing session and a single VP (Xu et al., 2017). The authors reported the inter-trial stability of the group amplitude obtained by averaging multiple bilateral electrodes. However, their report provided no information regarding the topographical reliability. Additionally, the investigated within-session reliability of trials spaced in the scale of minutes does not address the issue of reliability across longer temporal intervals.

In the present study, we sought to determine the reliability of the face discrimination responses (as measured with the FPVS paradigm) across two acquisition sessions separated by a 6-month interval. Aiming for a comprehensive understanding of individual differences in the face discrimination response, we computed reliability on three different aspects of the response. First, we addressed the reliability of the neural information-dependent response patterns observed across VPs. Second, we determined the reliability of the response amplitude independently for each VP, to assess the consistency of the response strength of individual observers. Third and finally, we computed each observer's VP-specific topographical reliability.

Collectively, our analyses aim to provide a comprehensive characterization of individual differences in neural face discrimination measured with the FPVS paradigm. We believe that this is a crucial prerequisite for studies aiming to deploy this paradigm for empirical and

clinical purposes, and to establish its functional significance in both healthy and impaired populations (Liu-Shuang et al., 2016; Xu et al., 2017).

## 2. Methods

### 2.1. Participants

We tested fourteen adults (seven females, one left-handed; mean age:  $24.5 \pm 3.8$  years) with normal or corrected-to-normal vision and no history of psychiatric or neurological disorders. Participants were either undergraduate students from the University of Fribourg, or acquaintances of the experimenters. All subjects provided written informed consent and received financial compensation for participation. The experiment was approved by the local ethics committee.

### 2.2. Stimuli and procedure

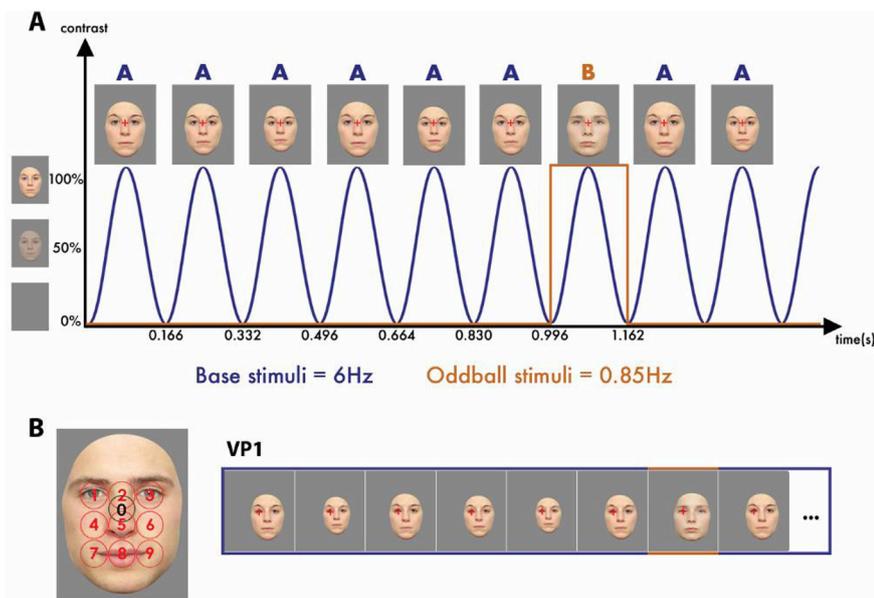
Stimulus material and procedures were identical to those previously reported (Liu-Shuang et al., 2014). Stimuli comprised full-front, colored images of 50 face identities (25 females) displaying neutral expressions. Images were cropped to exclude external facial features and were embedded in a grey background. Stimuli were presented on a VIEW-Pixx/3D monitor ( $1920 \times 1080$  pixel resolution, 120 Hz refresh rate) and subtended an average of  $11.02^\circ$  (height)  $\times$   $8.81^\circ$  (width) of visual angle at a viewing distance of 70 cm. In each 62s trial, face images were shown through sinusoidal contrast modulation at a frequency of 6 Hz, with image size varying randomly between 80 and 120% at each cycle to minimize pixel overlap (visual angle ranged between  $8.82$  and  $13.22^\circ$  (height) and  $7.05$ – $10.57^\circ$  (width)). A randomly selected face identity (base stimulus) was repeated throughout the sequence, with different face identities (oddball stimuli) interleaved every 7th face (i.e.  $6 \text{ Hz}/7 = 0.85 \text{ Hz}$ ; Fig. 1A). Each trial began and ended with 2s during which maximal image contrast progressively ramped up (fade in) and down (fade out), respectively. Observers were instructed to fixate a cross in the center of the screen while responding via button press to a brief (200 ms) change in the color of the fixation cross, which occurred randomly eight times within each trial. This served to ensure maintenance of fixation, and a constant level of attention. Subjects were accurate at this task ( $0.93 \pm 0.1$ ). Performance of one subject (S14) could not be computed during the first session due to technical issues.

We manipulated the viewing position (VP) by spatially arranging the stimuli in order to have one of ten facial regions behind the central fixation cross (Fig. 1B). Thus, these ten conditions differed in terms of the information being fixated by the observers throughout a trial. Note that VP0 was located slightly below the nasion, at what has been reported to be the “optimal” fixation position for old/new decisions under highly controlled viewing conditions at the group level (Hsaio & Cottrell, 2008). Each observer completed 20 trials (two repetitions per VP condition) during an initial acquisition session, and 20 trials during a second acquisition session 6-months later. Generally, each session comprised two runs of ten trials each. However, subjects could request breaks as required.

Both acquisition sessions involved identical stimuli and procedures; behavior was monitored through a webcam located inside the booth. In total, we obtained 40 trials (4 repetitions  $\times$  10 conditions) for each observer.

### 2.3. EEG acquisition

Electrophysiological data were acquired through the Biosemi ActiView software with a Biosemi Active-Two amplifier system recording from 128 Ag/AgCl active electrodes at a sampling rate of 1024 Hz. Four electrodes located at the outer canthi and below the eyes were placed for blink monitoring. In this active electrode system, the quality of electrode contact with the skin was evaluated by the offset relative to the magni-



**Fig. 1. Schematic illustration of the FPVS paradigm and the experimental conditions (A).** Faces were presented through sinusoidal contrast modulation. Base stimuli displayed the same facial identity and appeared at a frequency of 6 Hz. Different identity (oddball) faces were interleaved every 7th stimulus. Displayed here are stimuli presented at VP2. (B) In each condition faces were aligned in order to have one of the ten viewing positions (VPs) in the center of the screen and fixated by the subject (see right for an example).

tude of the feedback loop formed by the CMS-DRL electrodes, which was held below  $\pm 25$  mV throughout the recording. During data acquisition a 0.16–100 Hz filter was applied to remove any slow drift over time. Digital triggers were sent at the start and end of each stimulation sequence by the VIEWPixx screen to the acquisition computer.

#### 2.4. Analysis

**Preprocessing.** EEG data were processed with Letswave 5 (Mouraux and Iannetti, 2008). Continuous data were first digitally bandpass filtered to exclude frequencies below 0.1 Hz and above 100 Hz (4th order Butterworth filter). The signal was then downsampled to 256 Hz and segmented relative to each condition. For each observer, we extracted  $20 \times 66$ s epochs, which included 2 extra seconds pre- and post-stimulation. An independent component analysis using a square mixing matrix algorithm was computed to filter out noise related to blinks expressed by each subject (one single component was selected based on its topography and the correspondence between its temporal waveform and that of the ocular channels). Data were then visually inspected for interpolation of noisy electrodes (max. 5% of all scalp electrodes per observer were replaced through linear interpolation of the 3 surrounding electrodes). Data were then re-referenced to a common average reference and cropped to an integer number of oddball's cycles starting 2 s after stimulation onset and ending 2 s before stimulation offset (= 14932 bins).

**Frequency domain analysis.** The amplitude of EEG responses in the frequency-domain was extracted using Matlab's built-in Fast Fourier Transform (FFT) function. Baseline-correction was applied to all of the resulting amplitude spectra by subtracting from each frequency bin the average of its surrounding 20 bins (excluding the 2 immediately neighboring bins). For visualization purposes, we also converted the raw amplitude values into signal-to-noise ratio (SNR) by dividing each frequency bin by the average of its surrounding 20 bins (excluding the 2 immediately neighboring bins). We averaged the resulting signal over 24 occipito-temporal electrodes (A9-A16, A22-A29, B6-B11, D31-D32; cf., Fig. 2A) to include channels sensitive to both the general and the face discrimination response. Since the periodic response to our stimulation is spread over multiple harmonics (Fig. 2) (i.e. integer multiples of the stimulation frequency), we first determined the relevant range of frequency harmonics to take into account, independently for base and oddball frequencies. To this end, we z-scored the signal after averaging across conditions, subjects, and electrodes (Z-scores were computed

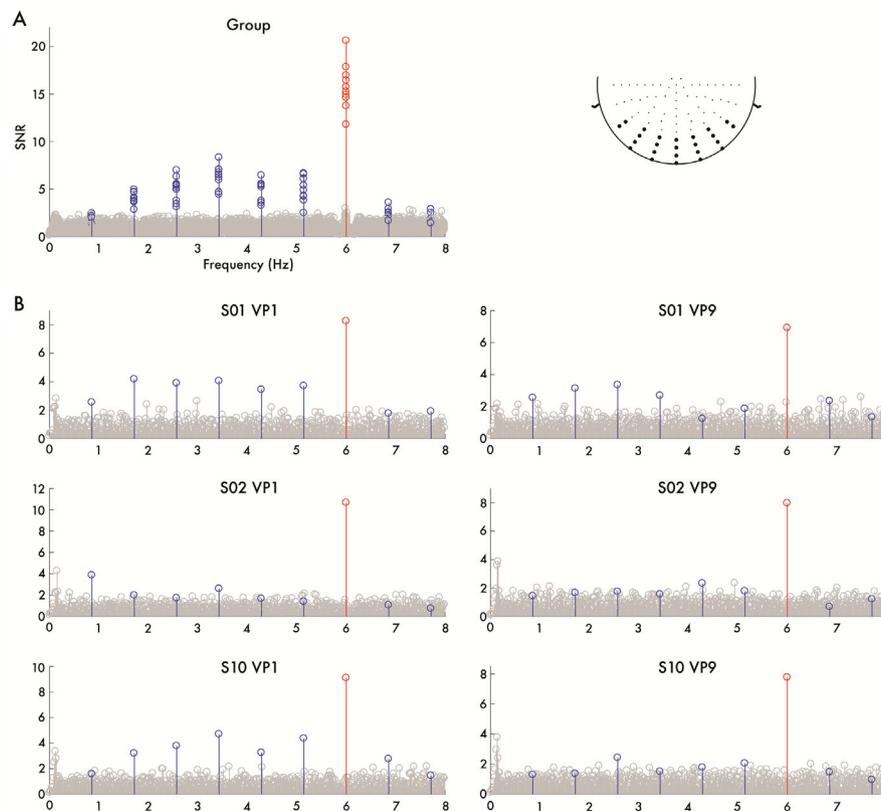
following the same logic as the baseline-correction). Harmonics were considered as significant until the z-scores of two consecutive harmonics failed to exceed 1.64 ( $p .05$ , one-tailed). Significant responses at the oddball frequency (0.85 Hz) and its harmonics reflect face identity discrimination, while responses at the base frequency (6 Hz) represent a combination of face-related processing and general visual responses to the stimulus presentation. Based on this threshold the oddball response, which indexes implicit neural face discrimination, was quantified by summing the first 9 oddball harmonics (i.e., 0.85–7.71 Hz), excluding the 7th harmonic since it is confounded with the base stimulation frequency rate. The base response was significant until the 9th harmonic (i.e., 6–54 Hz). However, since we were mainly interested in facial identity processing, and not the general visual response, we only considered the fundamental base frequency (i.e., 6 Hz) as a sanity-check for our experimental manipulation of fixated visual input.

#### 2.5. Statistical analysis

**Reliability of individual differences in VP-dependent patterns of response amplitude at the oddball frequencies.** To assess whether the patterns of response amplitude across VPs were stable across the two sessions, we carried out a regression between the two measures. We averaged trials from the same session and then regressed the face discrimination responses of session one (*AmplitudeSession1*) onto those of session two (*AmplitudeSession2*), using *Subject* as a random effect to account for individual differences (1). To avoid any a priori assumption regarding the topography of the effect, we performed the analysis for each scalp electrode independently, applying Holm-Bonferroni-correction to account for multiple-comparisons.

$$\text{AmplitudeSession1} \sim \text{AmplitudeSession2} + (1|\text{Subject}) \quad (1)$$

**Reliability of inter-individual differences in face discrimination response amplitudes.** To investigate whether inter-individual variations in response amplitude were stable over time, we computed their reliability using Cronbach's alpha coefficient. This measure, which has been used previously in a similar context (Xu et al., 2017), evaluates the similarity of responses across measurements. We computed Cronbach's alpha across four trials, for each VP and electrode independently. High reliability indicates that observers, who for example exhibit strong responses at one time point, are consistent across different trials and sessions. Conforming with standard practice, we chose a coefficient threshold of .8, which indicates that noise biases the responses by 20%, while the remaining



**Fig. 2. Frequency spectra.** (A) SNR spectrum averaged across 24 posterior electrodes (right) and all subjects, with all VP conditions overlapping (each dot represents the response magnitude for a VP condition). Orange peaks indicate the general visual response while blue peaks index the face identity discrimination response. (B) SNR spectrum averaged across the same 24 electrodes in three exemplary subjects for two VPs. Note that an SNR of 1 represents noise level.

80% represents the measure of interest (Nunnally, 1978). In addition to single-electrode analyses, to account for the variability across observers in the oddball response topography, we pooled 12 bilateral occipito-temporal electrodes and recomputed the reliability at each VP.

**Reliability of individual face discrimination response topographies.** To determine the stability of the scalp distribution of the response over time, topographical reliability was computed by means of Cronbach's alpha for each observer, at each VP independently. Signals that were subject to analysis were taken from posterior electrodes (cf., Fig. 6A), which included typically responsive channels in the context of neural face discrimination. Compared to amplitude computation, here we opted to consider a larger region of interest (ROI) in order to determine whether initially responsive clusters would be stable or change (expand, shift, or shrink) over time. High reliability of this posterior ROI would indicate that the electrodes showing stronger responses and forming responsive clusters are comparable across time. Conversely, low reliability would suggest that the response predominantly emerges at different electrodes at different time-points.

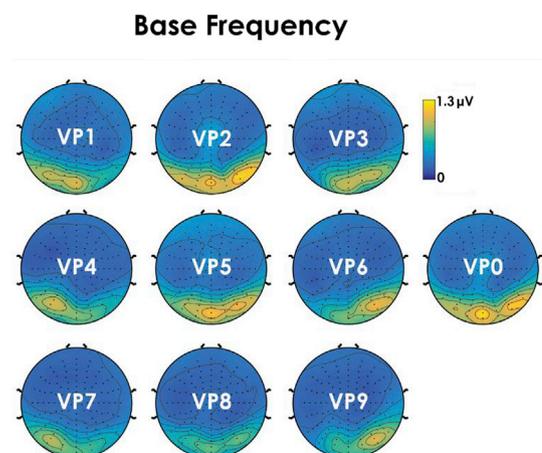
**Face discrimination response amplitude and topographical reliability.** We investigated the relationship between the topographical reliability (across 44 individual channels) and magnitude (average of a 12-channel ROI) of the face discrimination response using Holm-Bonferroni corrected Spearman correlations, independently for each VP.

### 3. Results

#### 3.1. Base frequency: general visual response

A response at the base frequency (6 Hz) indicates successful synchronization of the visual system to the contrast between background and face-stimuli (Dzhelyova et al, 2016; see Fig. 2). At the group level this general visual response was expressed predominantly at

occipito-temporal electrodes (Fig. 3). Conditions involving fixations on the left or right side of the face led to the majority of facial information being represented on the opposite visual field, and a shift of the response to the contralateral hemisphere. For example, fixation on the left eye (Fig. 3; VP1) led to most facial information being presented in the right visual field, which elicited strongest responses in the left hemisphere. Overall, averaging across 15 occipito-temporal electrodes (A10-A15, A23-A28, B7-B9) led to strongest general visual responses for VP2 (i.e., nasion VP; 1.05  $\mu$ V), followed by VP0 (1.04  $\mu$ V). Given this study's focus on the FPVS face identity discrimination response, the general visual responses expressed at the base frequency are not considered for further analyses.



**Fig. 3.** Grand average of the general visual response for the 10 VPs.

### 3.2. Oddball frequency: face discrimination response

Significant responses at the oddball frequency (i.e. 0.85 Hz = rate of face identity change) and its harmonics at occipito-temporal electrodes were found in all participants, and indicate successful neural face discrimination (Fig. 2B for three exemplary subjects) (Liu-Shuang et al., 2014). Here, we manipulated facial information fixated to determine the effect on subjects' neural face discrimination response, both in terms of amplitude and topography, and compare the inter-, and intra-individual differences across a time-interval of 6 months to estimate their reliability.

**VP-modulation of the face discrimination response.** We averaged the baseline-corrected signal from 4 trials to assess the effect of VP on the amplitude of the face discrimination response. In line with previous work (Liu-Shuang et al., 2016; Xu et al., 2017), this response was generally expressed at occipito-temporal electrodes (Fig. 4; Supplementary Fig. S1). Therefore, we quantified the neural face discrimination responses within a single bilateral occipito-temporal ROI comprising 12 electrodes (D31, D32, A9-A12, B6-B11) across which responses were averaged. This was done both at the group level (on the grand-average signal) and at the individual level.

At the group level, the face discrimination response amplitude was strongest for fixation at VP5 (i.e., nose; 1.33  $\mu$ V), followed by VP1 (i.e., left eye; 1.32  $\mu$ V) and VP0 (i.e., nasion; 1.25  $\mu$ V) (Fig. 4; for group and individual subject values see Supplementary Table S1). However, at the individual level, we observed strong inter-individual differences in the observer-specific response patterns, which deviated from the group level observation (Fig. 4; Supplementary Fig. S1). Contrary to the group, VP1 elicited the overall strongest response across observers (i.e., left eye, 4 subjects), followed by VP5 (i.e., nose, 3 subjects) (cf. Supplementary Table S1 for all values).

Mirroring these observations, individual observers' response topographies differed from the group topography. On average the response was right-lateralized for the majority of VPs. At the individual level, however, patterns were more heterogeneous and oddball responses could be located at right-, left-, or bilateral occipito-temporal electrodes (Fig. 4).

**Reliability of individual differences in VP-dependent patterns of response amplitude at the oddball frequencies.** To determine whether patterns of response amplitude across VPs observed at the first session would be predictive of those recorded 6 months later, we computed a linear regression between the two measures independently at each electrode. This analysis revealed a positive relationship at bilateral occipito-

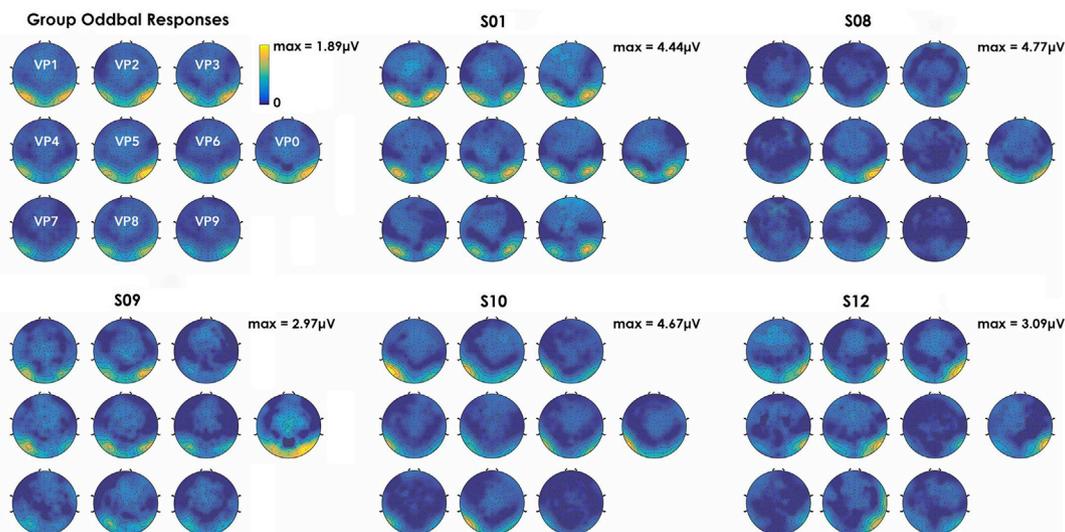
temporal clusters (Fig. 5A). The effect ranged from .39 (A10;  $p$  .001) to .52 (A28;  $p$  .001).

**Reliability of inter-individual differences in face discrimination response amplitudes.** Cronbach's alpha was computed across four trials for each VP and electrode independently to assess the reliability of inter-individual differences in amplitude. The results show that the most reliable responses were located at occipito-temporal electrodes. All VPs, with the exception of VP6 (right cheek), elicited reliable responses at least one channel (Fig. 5B). Overall, VP0 was the condition with the largest cluster of reliable electrodes. When averaging 12 bilateral occipito-temporal electrodes, reliability ranged from .89 (VP0) to .56 (VP6) (Fig. 5C). Contrary to analyses on single electrodes, here only five of 10 VPs elicited reliable responses in the selected ROI. Interestingly, these five VPs all involved fixation of the midline, or the left side of the face.

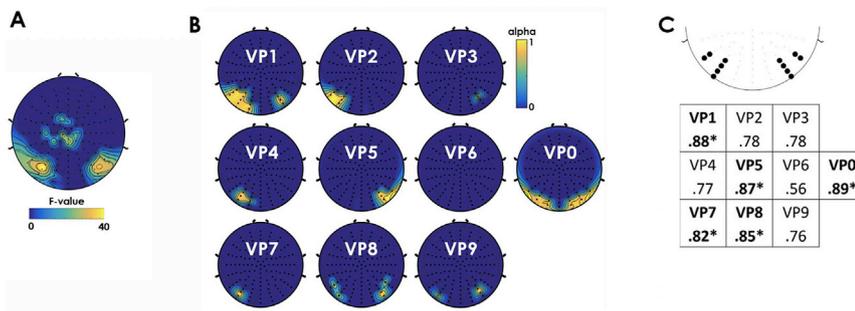
**Reliability of individual face discrimination response topographies.** We have shown that the response patterns across VPs, and the response amplitude within each VP were both reliable over a 6-month interval at the group level. To better understand potential individual differences in response reliability (i.e., Are all observers equally reliable in their response topography? Is each subject equally reliable across VPs?), we computed topographical reliability within each observer and independently for each VP.

Response consistency varied across VPs and observers. Across VPs, the average reliability pooled across all observers ranged from .62 (VP2; i.e., nasion) to .34 (VP9; i.e., right corner of the mouth). Across observers, the averaged alpha values from all VPs ranged from .88 (S10) to .09 (S07) (Fig. 6A). Across VPs and the entire cohort, the majority of observers (10/14) showed a reliable response topography for at least one VP. Two observers showed high topographical reliability for nearly all VPs (9/10 VPs), while eight exhibited reliable response topographies for only certain VPs (cf., S10 and S09, see Fig. 6). The remaining observers did not show any reliable responses between sessions (e.g., S07, see Fig. 6A).

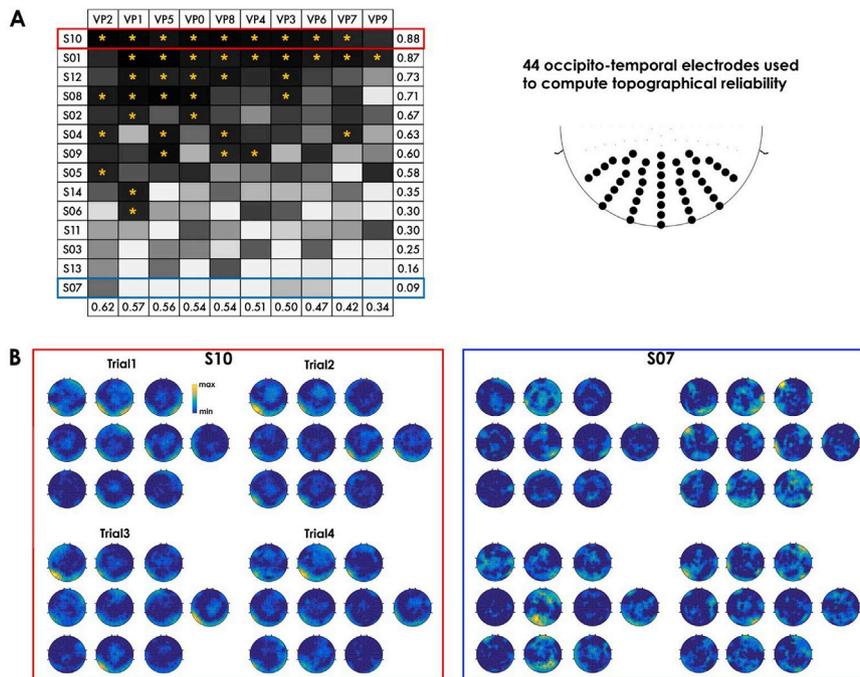
**Relationship between FPVS response amplitude and topographical reliability.** Visual exploration of our data suggested that subjects with more reliable topographies also exhibited stronger and more typical neural face discrimination response distributions over bilateral occipito-temporal clusters of electrodes. As illustrated in Fig. 7, correlation analyses showed significant positive relationships between the response amplitude pooled within this region, and the topographical reliability at each VP. The highest and lowest correlations were found for VP1 (left eye;  $r$  = .92,  $p$  .001) and VP6 (right cheek;  $r$  = .57,  $p$  = .03), respectively.



**Fig. 4. Neural face discrimination responses.** Face discrimination responses (quantified as the summed baseline-corrected amplitudes at oddball frequencies) are shown at the group-level and for five example subjects across the ten VP conditions. Though globally responses occur consistently over occipito-temporal channels, variation in response amplitude and topography are visible between and within subjects.



**Fig. 5. Reliability of the VP-patterns and FPVS response amplitude.** (A) Data-driven regression results. F-values of non-significant electrodes were set to 0. VP-patterns were most stable at occipito-temporal electrodes. (B) Reliability of response amplitude across subjects is shown for each condition and each electrode respectively. Alpha values below the .8 threshold were set to 0 for visualization purposes. (C) Reliability of response amplitude computed across 12 occipito-temporal electrodes (top) for each VP.



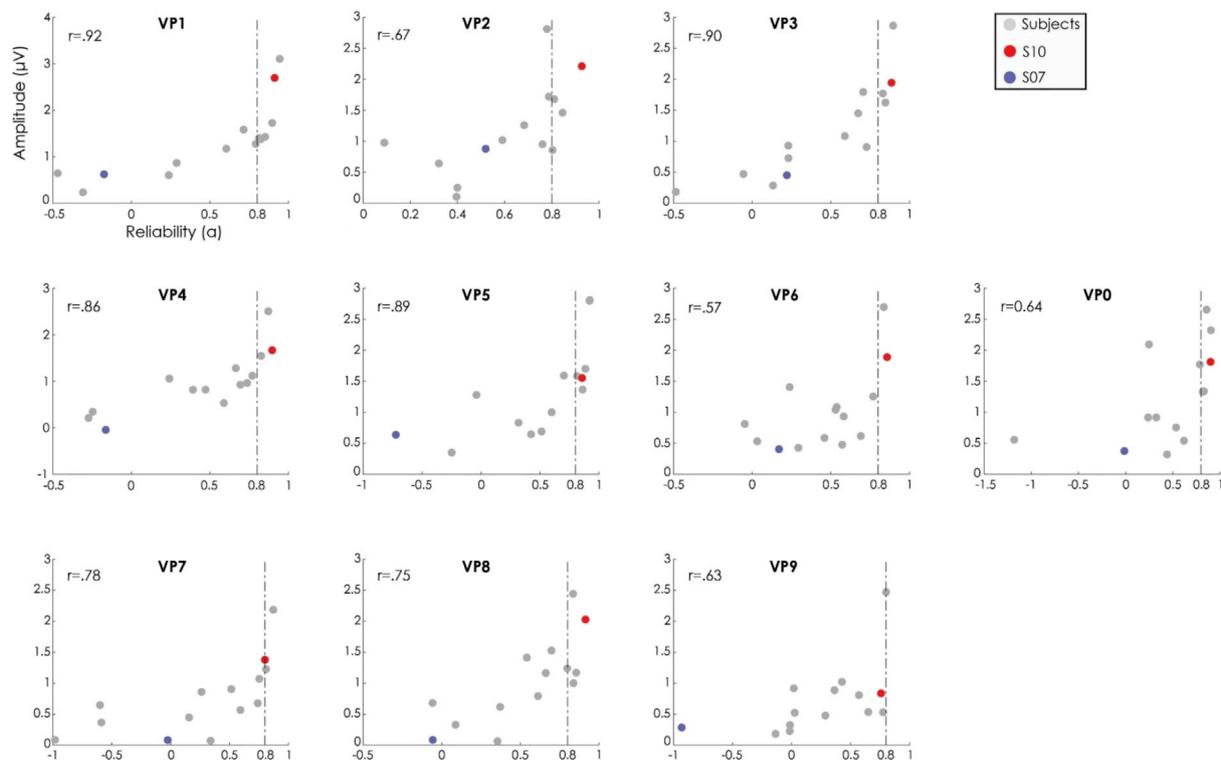
**Fig. 6. Topographical reliability of the FPVS response.** (A) Individual topographical reliability computed across 44 occipito-temporal electrodes for each subject and VP. The color scale ranged between alpha = 0 (white) and 1 (black). Asterisks indicate alpha  $\geq$  .8. (B) VP-dependent FPVS responses for the most (left) and least (right) reliable subject across all four trials.

#### 4. Discussion

The present study aimed to provide novel insights into individual differences in FPVS face identity discrimination recorded by EEG. We explored inter-individual variability by considering the influence of facial information being fixated (viewing position; VP), and the consistency of face discrimination responses over time. Our data underline the effect of VP on responses not only at the group level, but especially at the level of individual observers, who exhibited idiosyncratic patterns of responses across VPs. Specifically, for any given VP, we observed individual differences – both in terms of response amplitude, as well as its topography. Our reliability analysis showed that the VP patterns recorded at the first session were largely predictive of those observed at the second session, i.e. 6 months later. Additionally, individual differences in the stability of response amplitude systematically occurred across trials and sessions for the majority of the VPs at occipito-temporal electrodes. Concerning topographical reliability, we report a considerable degree of variation of the scalp distribution of the neural face discrimination responses across observers. Finally, we observed a positive relationship between the topographical reliability of individual observers’ responses, and the response amplitude over occipito-temporal electrodes.

Recently, Xu et al. (2017) reported individual differences in the amplitude EEG face discrimination responses indexed by fast periodic visual stimulation (FPVS). Similar to numerous neuroimaging (and behavioral) studies, the authors standardized the visual input by instructing subjects to fixate on one predefined VP (VP0, i.e. the center of the face). In the present study we observed that subjects exhibited idiosyncratic patterns of responses across VPs: while some subjects show stronger responses when fixating central VPs, others exhibit higher responses to facial information in other regions (e.g., VP1, left eye). Our findings indicate that observations made based on a single location (i.e. standardized input) may lead to a misrepresentation of a large proportion of subjects. The importance of these idiosyncrasies is further highlighted by the fact that the patterns observed at one session could predict those exhibited 6 months later. This suggests that the robust “neural preference” for some VPs over others is unlikely to be noise-related. Instead, the observed reliable VP-dependent responses at occipito-temporal sites are likely to reflect intrinsic characteristics of each individual’s face-sensitive neural networks.

Together, these observations may reconcile previous seemingly inconsistent findings regarding the facial features eliciting the strongest face-sensitive neural response. The majority of studies addressing this



**Fig. 7.** Scatter plots demonstrating the relationship between individual observers' response amplitude and response reliability. For each VP, the average amplitude across trials over occipito-temporal channels is plotted against the reliability score across individual trials for all subjects. Stronger face discrimination responses were more reliable. The dotted line marks the threshold for a response to be considered reliable ( $\alpha = 0.8$ ).

question reported that fixation on the eyes leads to stronger responses compared to other features (e.g., Nemrodov et al., 2014; de Lissa et al., 2014). However, other studies suggest that fixation on other facial regions evokes comparable or stronger responses (e.g., Zerouali et al., 2013; McPartland et al., 2010). Importantly, these different findings were made in the context of group studies, which stressed commonalities (at the expense of individual differences) in neural responses. Here, adopting an individual observer approach, we observed that while many subjects exhibited stronger responses for the eye-region, others exhibited “neural preferences” for VPs closer to the center of the face (i.e., VP2/nasion, VP0/center, VP5/nose). This variation in the population questions the existence of a *per se* optimal position for face processing (see Caldara, 2017), and could account for discrepant findings reported in the context of earlier group-studies.

Our data also show that individual differences modulate a given measure – here the FPVS face discrimination response – on different levels. In the present context, individual differences could be expressed in terms of the neural response amplitude, scalp distribution, or response reliability. This last parameter is crucial to determine whether an observation carries potentially meaningful information, as opposed to reflecting random noise. Altogether, our results lead to two main considerations.

First, although individual differences can provide a unique source of information, not all should be considered equally meaningful. Disregarding whether an individual exhibits reliable responses over time can mask effects in a similar vein as does averaging across observers. In order to identify *genuine* variations of interest, reliable responses have to be separated from unreliable ones, which inherently carry more noise. This is important in studies investigating the relationship between multiple responses. Those obtained from unreliable observers are less likely to correlate across measures and will therefore introduce noise, potentially masking otherwise stronger relationships. Similarly, quantifying individuals' reliability would also facilitate determining commonalities across subjects.

Second, individual differences in reliability also indicate that this parameter should ideally be assessed for any tested cohort. More concretely, an estimate of good internal consistency reported for a measure within one group of observers cannot be expected to exist in the context of other independent observers. Individual subjects' reliability is an additional parameter that could be reported along with a given finding to help its interpretation and assess the extent to which it can be generalized. We suggest that – at least in the context of face discrimination measured with FPVS – response amplitude over occipito-temporal regions can provide an initial estimate of its stability over time.

The relationship we report here suggests that precise and stable topographies may relate to more efficient neural responses. This hypothesis is in line with recent findings reporting that the magnitude of face sensitive neural responses were associated with behavioral face recognition (Xu et al., 2017; Elbich and Scherf, 2017). Therefore, functionally meaningful topographical reliability may represent an additional promising parameter to track e.g. development changes in face processing, or effects of interventions in clinical populations. A covariance between neural response reliability and amplitude at the individual level could also be exploited as an objective index to track potential improvements or impairments in face processing. Further research is required to understand the cause(s) of differential neural response reliability – functional or anatomical – as well as the directionality of the observed relationship between reliability and response magnitude.

## Acknowledgements

This work was supported in part by an FSR-FNRS postdoctoral grant awarded to JLS [grant no: FC 91608].

## Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.neuroimage.2019.01.023>.

## References

- Arizpe, J., Walsh, V., Yovel, G., Baker, C.I., 2017. The categories, frequencies, and stability of idiosyncratic eye-movement patterns to faces. *Vis. Res.* 141, 191–203.
- Blais, C., Jack, R.E., Scheepers, C., Fiset, D., Caldara, R., 2008. Culture shapes how we look at faces. *PLoS One* 3 (8), e3022.
- Caldara, R., 2017. Culture reveals a flexible system for face processing. *Curr. Dir. Psychol. Sci.* 26 (3), 249–255.
- Das, K., Giesbrecht, B., Eckstein, M.P., 2010. Predicting variations of perceptual performance across individuals from neural activity using pattern classifiers. *Neuroimage* 51 (4), 1425–1437.
- de Lissa, P., McArthur, G., Hawelka, S., Palermo, R., Mahajan, Y., Hutzler, F., 2014. Fixation location on upright and inverted faces modulates the N170. *Neuropsychologia* 57, 1–11.
- Dzhelyova, M., Jacques, C., Rossion, B., 2016. At a single glance: fast periodic visual stimulation uncovers the spatio-temporal dynamics of brief facial expression changes in the human brain. *Cerebr. Cortex* 27 (8), 4106–4123.
- Elbich, D.B., Scherf, S., 2017. Beyond the FFA: brain-behavior correspondences in face recognition abilities. *Neuroimage* 147, 409–422.
- Gaspar, C.M., Rousselet, G.A., Pernet, C.R., 2011. Reliability of ERP and single-trial analyses. *Neuroimage* 58 (2), 620–629.
- Henderson, J.M., Williams, C.C., Falk, R.J., 2005. Eye movements are functional during face learning. *Mem. Cognit.* 33 (1), 98–106.
- Herzmann, G., Kunina, O., Sommer, W., Wilhelm, O., 2010. Individual differences in face cognition: brain-behavior relationships. *J. Cognit. Neurosci.* 22 (3), 571–589.
- Hsiao, J.H., Cottrell, G., 2008. Two fixations suffice in face recognition. *Psychol. Sci.* 19 (10), 998–1006.
- Kaltwasser, L., Hildebrandt, A., Recio, G., Wilhelm, O., Sommer, W., 2014. Neurocognitive mechanisms of individual differences in face cognition: a replication and extension. *Cognit. Affect Behav. Neurosci.* 14 (2), 861–878.
- Kanan, C., Bseiso, D.N.F., Ray, N.A., Hsiao, J.H., Cottrell, G.W., 2015. Humans have idiosyncratic and task-specific scanpaths for judging faces. *Vis. Res.* 108, 67–76.
- 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.
- Liu-Shuang, J., Torfs, K., Rossion, B., 2016. An objective electrophysiological marker of face individualisation impairment in acquired prosopagnosia with fast periodic visual stimulation. *Neuropsychologia* 83, 100–113.
- McPartland, J., Cheung, C.H., Perszyk, D., Mayes, L.C., 2010. Face-related ERPs are modulated by point of gaze. *Neuropsychologia* 48 (12), 3657–3660.
- Mehouard, E., Arizpe, J., Baker, C.I., Yovel, G., 2014. Faces in the eye of the beholder: unique and stable eye scanning patterns of individual observers. *J. Vis.* 14 (7), 6.
- Mouraux, A., Iannetti, G.D., 2008. Across-trial averaging of event-related EEG responses and beyond. *Magn. Reson. Imag.* 26 (7), 1041–1054.
- Nemrodov, D., Anderson, T., Preston, F.F., Itier, R.J., 2014. Early sensitivity for eyes within faces: a new neuronal account of holistic and featural processing. *Neuroimage* 97, 81–94.
- Norcia, A.M., Appelbaum, L.G., Ales, J.M., Cottoreau, B.R., Rossion, B., 2015. The steady-state visual evoked potential in vision research: a review. *J. Vis.* 15 (6), 4.
- Nunnally, J.C., 1978. *Psychometric Theory*, second ed. McGraw-Hill, New York.
- Peterson, M.F., Eckstein, M.P., 2013. Individual differences in eye movements during face identification reflect observer-specific optimal points of fixation. *Psychol. Sci.* 24 (7), 1216–1225.
- Schyns, P.G., Bonnar, L., Gosselin, F., 2002. Show me the features! Understanding recognition from the use of visual information. *Psychol. Sci.* 13 (5), 402–409.
- Seghier, M.L., Price, C.J., 2018. Interpreting and utilising intersubject variability in brain function. *Trends Cognit. Sci.* 22 (6), 517–530.
- Turano, M.T., Marzi, T., Viggiano, M.P., 2016. Individual differences in face processing captured by ERPs. *Int. J. Psychophysiol.* 101, 1–8.
- Walker-Smith, G.J., Gale, A.G., Findlay, J.M., 1977. Eye movement strategies involved in face perception. *Perception* 6 (3), 313–326.
- Wilhelm, O., Herzmann, G., Kunina, O., Danthiir, V., Schacht, A., Sommer, W., 2010. Individual differences in perceiving and recognizing faces — one element of social cognition. *J. Pers. Soc. Psychol.* 99 (3), 530.
- Xu, B., Liu-Shuang, J., Rossion, B., Tanaka, J., 2017. Individual differences in face identity processing with fast periodic visual stimulation. *J. Cognit. Neurosci.* 29 (8), 1368–1377.
- Yarbus, A.L., 1967. Eye movements during perception of complex objects. In: *Eye Movements and Vision*. Springer US, Boston, MA, pp. 171–211.
- Zerouali, Y., Lina, J.M., Jemel, B., 2013. Optimal eye-gaze fixation position for face-related neural responses. *PLoS One* 8 (6), e60128.