

# **Varying stimulus duration reveals consistent neural activity and behavior for human face individuation**

Talia L. Retter<sup>1,2,3</sup>, Fang Jiang<sup>2</sup>, Michael A. Webster<sup>2</sup>, Caroline Michel<sup>1</sup>,  
Christine Schiltz<sup>3</sup> & Bruno Rossion<sup>1,4,5</sup>

<sup>1</sup> Psychological Sciences Research Institute, Institute of Neuroscience,  
UCLouvain (Belgium)

<sup>2</sup> Department of Psychology, Center for Integrative Neuroscience, University of  
Nevada, Reno (USA)

<sup>3</sup> Department of Behavioural and Cognitive Sciences, Institute of Cognitive Science &  
Assessment, University of Luxembourg (Luxembourg)

<sup>4</sup> Université de Lorraine, CNRS, CRAN, F-54000, Nancy, France

<sup>5</sup> Université de Lorraine, CHRU-Nancy, Service de Neurologie, F-54000 Nancy,  
France

Corresponding author:

Talia L. Retter

University of Luxembourg

Maison des Sciences Humaines

11, Porte des Sciences

4365 Esch-sur-Alzette

Luxembourg

+352 46 66 44 6497

talia.retter@uni.lu

## Abstract

Establishing consistent relationships between neural activity and behavior is a challenge in human cognitive neuroscience research. We addressed this issue using variable time constraints in an oddball frequency-sweep design for visual discrimination of complex images (face exemplars). Sixteen participants viewed sequences of ascending presentation durations, from 25 to 333 ms (40 to 3 Hz stimulation rate) while their electroencephalogram (EEG) was recorded. Throughout each sequence, the same unfamiliar face picture was repeated with variable size and luminance changes while different unfamiliar facial identities appeared every 1 s (1 Hz). A neural face individuation response, tagged at 1 Hz and its unique harmonics, emerged over the occipito-temporal cortex at 50 ms stimulus duration (25 to 100 ms across individuals), with an optimal response reached at 170 ms stimulus duration. In a subsequent experiment, identity changes appeared non-periodically within fixed-frequency sequences while the same participants performed an explicit face individuation task. The behavioral face individuation response also emerged at 50 ms presentation time, and behavioral accuracy correlated with individual participants' neural response amplitude in a weighted middle stimulus duration range (50 to 125 ms). Moreover, the latency of the neural response peaking between 180 to 200 ms correlated strongly with individuals' behavioral accuracy in this middle duration range, as measured independently. These observations point to the minimal (50 ms) and optimal (170 ms) stimulus durations for human face individuation and provide novel evidence that inter-individual differences in the magnitude and latency of early, high-level neural responses are predictive of behavioral differences in performance at this function.

## Introduction

Within tens of milliseconds, the human brain makes sense of complex visual inputs from the environment. This time frame includes low-level, sensory visual processing, such as photoreceptor transduction and the active transformation of visual inputs from the retina to the lateral geniculate nucleus and then the primary visual cortex (Hubel & Wiesel, 2005; Gabbiani & Cox, 2017). Moreover, this time frame includes higher-level processing, in which a network of ventral occipito-temporal brain regions supports visual object recognition (Kravitz et al., 2013; Grill-Spector & Weiner, 2014). The rapid speed of human visual recognition enables interactions with spatially complex environments that are dynamically changing, even when dealing with discriminations of highly similar visual stimuli, as in the case for human faces.

In *face individuation* (FI), a perceiver discriminates individuals' faces from each other, as well as generalizes each individual's face across variable viewing conditions (Rossion, Retter & Liu-Shuang, 2020). To accomplish this function, a highly calibrated recognition system is required: the physical differences between human facial identities may be quite small, and the variations in viewing the same facial identity may be quite large, e.g., from changes in distance, lighting, viewpoint, etc. (Jenkins et al., 2011). Moreover, for comfortable social interactions, FI must be completed rapidly, and not only for learned, familiar faces: even when encountering an unfamiliar face, one must establish quickly that this face is unknown, and to be able to discriminate it from other individual faces in the environment (e.g., a crowd). However, how long a face should be viewed for above chance and maximal individuation performance is presently unclear.

Some studies reported maximal identification performance for half a dozen learned face images presented for as low as a 4 ms duration (Bachmann, 1991; Gur, 2018), but a lack of stimulus masking in these studies could have led to long post-stimulus processing. Using an old/new behavioral recognition task with pictures of 64 unfamiliar faces, Hsiao & Cottrell (2008) showed that on average 1.81 gaze fixations were sufficient for maximal performance (see also Orban de Xivry et al., 2008; Peterson & Eckstein, 2013); however, the average fixation duration lasted about 600 ms in that study, and stimulus duration was not manipulated below that value.

A number of studies did restrain image processing time, but used limited stimulus sets. For example, two studies that used backward masking produced very different results, perhaps due to stimulus differences across only 5-6 images. Rolls et al. (1994) used 5 photographs of highly familiar faces, with variations in gender and external cues, to report above-chance

identification with only 16 ms (masked after 4 additional ms; 20 ms stimulus-onset-asynchrony; SOA), with maximal performance reached for a 36 ms stimulus duration (40 ms SOA). In contrast, Tanskanen et al. (2007) used images of 6 familiarized male faces to report above-chance identification from 50 ms stimulus duration, with maximal recognition rates reached at 100 ms. In further comparison, blending 6 familiarized images of male faces with 15 distractors in rapid sequences of images suggested that about 100 ms was the threshold for reliable (about 80% accurate) FI (Nasanen et al., 2006). Finally, some studies used large but unnatural stimulus sets. In one such study, sensitivity at differentiating 4 sets of 41 subtly morphed, synthetic stimuli, with a delayed match-to-sample paradigm incorporating backward masking, was reported as maximal with 100 ms stimulus duration (Lehky, 2000). More recently, Or & Wilson (2010) used 81 synthetic face/anti-face stimulus pairs, defined by geometric information rather than facial features, and a two-alternative forced-choice task with backward masking, to report a threshold of 63-ms stimulus duration for individuation, with maximal performance apparent at about 107 ms.

Overall, while these behavioral studies provide useful information, they all rely on forced-choice matching or old/new recognition of (usually few) identical images, leading to short stimulus duration minima (as low as 4-16 ms) and ceilings (ranging from about 40-110 ms) for FI performance. Moreover, while there is a great deal of interest for inter-individual variability in cognitive functioning in the human population, including FI ability (with individual performance usually compared in terms of ability to individuate pictures of unfamiliar faces, e.g., Burton et al., 2010; Bowles et al., 2009; Hildebrandt et al., 2011; McCaffery et al., 2018; Rossion & Michel, 2018; see also Xu et al., 2017; Stacchi et al., 2019), whether, and to which extent, people vary in their ability to individuate complex visual images across variations of viewing time has not been addressed to our knowledge.

At the neural level, Tanskanen et al. (2007) tested their participants with magnetoencephalographic (MEG) recordings, although not with FI, but with faces *vs.* phase-scrambled stimuli. This showed two components (M170 and M300) emerging from 50 ms of face stimulus duration, in line with behavioral performance recorded in that study; however, contrary to behavior, amplitude of these components continued to increase until the longest duration of 200 ms, so that the optimal duration could not be defined precisely. In an EEG study by Alonso-Prieto et al. (2013), sequences of constantly varying (“different”) facial identities were contrasted to sequences of one repeating (“same”) facial identity across 14 presentation rates, from 1 to 16.7 Hz (i.e., 1000 to 60 ms SOA, with sinusoidally contrast-modulated image presentation). Larger EEG responses for different than same facial identity

conditions (i.e., a neural adaptation/repetition suppression effect; see Grill-Spector, Henson & Malach, 2006) were reported at a 120 ms duration (8.33 Hz) and above, with a maximal difference at 170 ms (5.88 Hz; this maximal difference was replicated with functional magnetic resonance imaging in Gentile & Rossion, 2014). These results could be taken as an indication that a minimum stimulus duration of about 120 ms is required for FI, except that the interference of sequential EEG responses at high rates confounds the effects of viewing time *per se* (e.g., Keyzers et al., 2001; Keyzers & Perrett, 2002; Retter et al., 2016; 2020).

This brief survey serves to illustrate that human FI performance, taken across stimulus presentation durations, has not yet been considered with both behavioral and neural measures in the same methodological framework. Therefore, their integrated interpretation is unknown. This lack of knowledge is surprising because the relationship between behavioral and neural FI responses has emerged as a topic of great interest in human face recognition research in recent years (e.g., Herzmann et al., 2010; Kaltwasser et al., 2014; Hermann et al., 2017; Xu et al., 2017; McGugin et al., 2018; Wu et al., 2018; Dobs et al., 2019; Dzhelyova et al., 2020). Unfortunately, several variables must be considered when assessing neural and behavioral responses (i.e., amplitude and latency of neural responses in space and time; accuracy and response time of behavioral responses) and inter-individual variability in both these measures can be both due to multiple general factors (see Rossion, Retter & Liu-Shuang, 2020), making this research program challenging. Perhaps for these reasons, whether large and early neural signatures of FI relate to better behavioral performance at this function in a neurotypical human population, for instance, is presently unknown.

Here, we address this question by investigating the impact of stimulus duration on FI, both behaviorally and neurally, with a focus on relating inter-individual differences at both measures. We used a novel combination of an oddball paradigm for measuring robust, high-level neural FI responses, even at the individual participant level (from Liu-Shuang et al., 2014; review: Rossion, Retter & Liu-Shuang, 2020) and a frequency-sweep design for progressively increasing stimulus duration within each stimulation sequence, in 11 steps from 25 to 333 ms (40 to 3 Hz), with forward- and backward-masking deriving from a 0 ms inter-stimulus-interval (from Retter et al., 2020; see Methods). We define the *minimal* stimulus duration for FI (i.e., the smallest stimulus duration to elicit a significant response), with convergent results produced for neural and behavioral measures, as well as the *optimal* stimulus duration (i.e., the smallest stimulus duration generating the largest response). Further, we investigate at which stimulus durations individual differences were most pronounced, and, taking stimulus duration into account, related differences in neural FI response amplitude with behavioral performance.

Finally, we probe neural FI responses in the time domain, to strengthen the characterization of the relationship of individuals' neural and behavioral responses under time constraints.

## **Experimental Procedures**

### **Participants**

Sixteen healthy, neurotypical adults participated in this experiment (age: mean = 23 years old; SD = 2.3 years; range = 19 to 28 years; gender identification: 10 female; 6 male; all right-handed (according to an adapted Edinburgh Handedness Inventory; Oldfield, 1971), all reporting normal or corrected-to-normal visual acuity. All were recruited from the UCLouvain community, and compensated monetarily for their time; none were excluded from the analyses. The Institutional Review Board of UCLouvain approved the study protocols, which were carried out in accordance with the Code of Ethics of the World Medical Association (Declaration of Helsinki).

### **Stimuli and Display**

Images of 100 different facial identities were used in this experiment (**Fig. 1**). All were derived from photographs of full-front, expressionless, female faces, taken under standardized conditions. They were minimally processed in Adobe Photoshop CS5: they were cropped around the border of the face to remove external features and resized to the same height, but the overall shape difference between faces was preserved. Note that a subset of these stimuli has been used in previous oddball EEG individuation studies (e.g., Liu-Shuang et al., 2014; Xu et al., 2017; Yan, Liu-Shuang & Rossion, 2019; first presented in a behavioral study by Laguesse et al., 2012). While other stimulus features were not standardized across facial identities (width, luminance, contrast, color, etc.), these aspects were modulated at every stimulus presentation, in order to increase their variability for each identity and thus reduce their diagnosticity across different identities (i.e., low-level stimulus control by variability, e.g., Thorpe, Fize & Marlot, 1996; Crouzet et al., 2010; Foldiak et al., 2004; Rossion et al., 2015; Retter et al., 2020). At every presentation, the stimulus size varied randomly from 80 to 120% of the original size (sampled in 5% steps; see Dzhelyova & Rossion, 2014), and the luminance varied randomly from -10 to +10% of the original (sampled in 2.5% steps). The stimuli were presented with a liquid-crystal display testing monitor, with a refresh rate of 120 Hz and a resolution of 1920 by 1080 pixels. Viewed from a distance of 80 cm, the stimuli subtended approximately 5.0 degrees of vertical visual angle at the original presentation size; width varied according to the individual face exemplars.



**Figure 1.** Stimuli: the 100 different female facial identities used in the study, all unfamiliar to the participants.

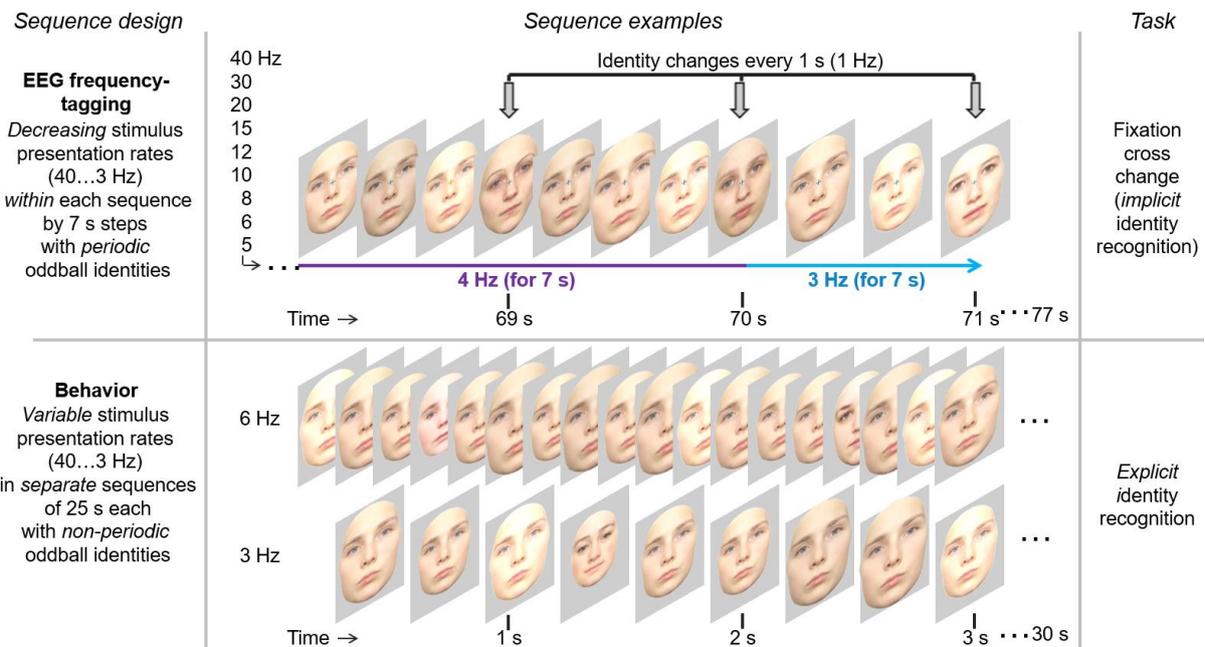
### EEG Frequency-Tagging Procedure

This experiment was based on a novel combination of two recently established EEG frequency-tagging approaches: 1) an oddball paradigm to measure high-level face individuation (FI; since Liu-Shuang et al., 2014; reviewed in Rossion, Retter & Liu-Shuang, 2020); and 2) a frequency-sweep design (Retter et al., 2020). In the oddball paradigm, one facial identity is repeated as the “base” face, while randomly selected “oddball” faces are interleaved at a fixed interval, i.e., as every  $n^{\text{th}}$  stimulus. There are thus two frequencies tagged: the stimulus presentation frequency ( $F$ ), and the oddball presentation frequency ( $F/n$ ). The stimulus presentation frequency,  $F$ , measures both low- and high-level visual responses common to the face stimuli presented; the oddball presentation frequency,  $F/n$ , is a differential response, reflecting the differences in the responses to base and oddball faces, i.e., measuring FI (for a review: Rossion, Retter & Liu-Shuang, 2020). Since  $F/n$  is a relatively low frequency (e.g., 1 Hz), FI EEG responses overlapping in time are avoided, and FI responses can be investigated over a long (1 s) window in the time domain. In the frequency-sweep design,  $F$  is continuously swept through descending frequency rates within each testing sequence (e.g., from 40 to 30 to 20 Hz (i.e., 25 to 33 to 50 ms), etc.), while  $F/n$  remains constant at 1 Hz (1 s; Retter et al., 2020). This design was created to measure the first, shortest duration (i.e., highest frequency) at which the differential  $F/n$  response would appear, as well as to characterize  $F/n$  (relative to  $F$ ) responses across stimulus presentation rates.

In the present combination of the oddball FI paradigm and the frequency-sweep design, each 77-s sequence contained one base facial identity throughout, despite changes in  $F$ , and the oddball identity appearing at  $F/n$  was randomly selected from the remaining 99 facial identities

at each oddball presentation, except that no oddball identities were repeated within a sequence. A large number of oddball facial identities were used to decrease the probability that FI responses would depend on (low-level) stimulus-specific features; however, the same base facial identity was used within a sequence to limit variability across stimulus presentation rates. Further, three different sets of 18 unique base identities were used across participants: each set was shown to six different participants.

Within each 77-s testing sequence,  $F$  was continuously swept from 40 Hz to 3 Hz (25 to 333 ms) in 11 steps of 7 s each, in accordance with the 120 Hz monitor. Specifically,  $F$  progressed as follows: 40 Hz, 30 Hz, 20 Hz, 15 Hz, 12 Hz, 10 Hz, 8 Hz, 6 Hz, 5 Hz, 4 Hz, and 3 Hz (25 ms, 33 ms, 50 ms, 67 ms, 83 ms, 100 ms, 125 ms, 167 ms, 200 ms, 250 ms, and 333 ms). Crucially, despite the changes in  $F$ , the FI frequency of  $F/n$  was held constant at 1 Hz (1 s) throughout each sequence (**Fig. 2**). To demonstrate the stimulation on a more common 60 Hz monitor, a movie was created with the nine available frequencies at 60 Hz (30, 20, 15, 12, 10, 6, 5, 4, and 3 Hz; i.e., 33, 50, 67, 83, 100, 167, 200, 250, and 333 ms; **Movie S1**). Images were presented continuously, at maximal contrast for the entire stimulus presentation duration, i.e., with a 0 ms inter-stimulus interval and a 100% duty cycle (Retter et al., 2018; 2020). There were 18 repetitions of this oddball frequency-sweep sequence, for a total of 126 s, and 126 oddball identity presentations, per condition.



**Figure 2.** Upper row: An overview of the stimulation sequences in the EEG frequency-tagging experiment, with each stimulation frequency defining a stimulus duration (e.g., 40 Hz = 25 ms) (see also **Movie S1**). Note the changes in size and luminance at every stimulus presentation,

*with the change of identity occurring every 1 s (1 Hz) throughout the sequence. Lower row: A couple example sequence segments from the behavioral experiment.*

Each trial consisted of: 1) a fixation cross presented in the center of the screen for 2-4 s, to establish attention; 2) the 77-s testing sequence as described above; 3) the fixation cross for another 2-4 s, to limit (eye) movements at the end of the sequence. Participants were instructed to fixate on the fixation cross, and to press on a keyboard key (space bar) each time the cross briefly changed color (blue to dark yellow, for 250 ms), which occurred at random intervals, 15 times per trial. Participants were not given any information about facial identity changes; there were only told that sequences of face images would be presented at different, decreasing speeds within each trial. Standardized testing conditions were maintained: the viewing distance from the testing monitor was measured at the beginning of the experiment, in a quiet room, dimly lit with a halogen lamp and the computer monitors, and the participant's behavior was monitored throughout by the experimenter recording the EEG. The total EEG recording session lasted about 35 minutes, including short rest breaks in between trials.

At the end of the experiment, participants were asked two questions: 1) an open-ended question about what they noticed during the experimental trials; and 2) whether they noticed any periodicity within the sequences. In response to the first question, most participants (75%) reported that one facial identity would repeat and sometimes there would be changes. Of these participants, some estimated the amount of the repeating face as: 40%, 1/3, 2/3, or "4/5, 2/3, or at least more than 1/2"; one participant commented that this occurred in some trials but not others. Other occasional observations included that only Caucasian, female faces were presented; that stimuli were presented at different rates, sometimes noting that the rate progressively slowed; that faces changed size; that there were different faces in a trial; that variations occurred in eye color and face contour; that some faces appeared deformed; and that the repeated face changed across trials. In response to the second question, no participants noticed any periodicity of the time of facial identity changes.

### **Behavioral Experiment Procedure**

The behavioral experiment was tested with the same participants in a separate testing session, following the EEG experiment (days between testing sessions: mean = 21 days; SD = 12.5 days). Behavior and EEG were recorded separately for two main reasons. First, the periodicity of facial identity changes was present in the EEG experiment (to enable implicit, frequency-tagged FI response analysis) but was not present in the behavioral experiment (so that explicitly detected facial identity changes did not occur at a predictable, fixed interval). Second, by not

having explicit behavioral responses to facial identity changes in the EEG experiment, we avoided contamination of the neural FI response with related decisional and motor brain processes. In the behavioral experiment, participants were first debriefed about the EEG experiment (as in Retter et al., 2020); in particular, they were told that identity changes had occurred exactly every 1 s. They were subsequently given instructions for the behavioral experiment, an explicit FI task. They were instructed to press on a keyboard key (J) with the index finger of their right hand each time they detected a different facial identity appearing in the sequence (i.e., an identity other than that of the base face), which could occur non-periodically, or not at all, in relatively short (25 s) sequences. Within a sequence, there would be only one stimulus presentation frequency ( $F$ ), but different sequences would have different presentation speeds.

Participants were told that the task could be very difficult, especially at high stimulus presentation speeds, but that they should try their best to respond as accurately as possible. They began with a demo trial, consisting of a base facial identity and presentation frequency (1.5 Hz; 667 ms) not used in the main experiment, to ensure that they understood the task. In actuality, different facial identities were inserted 4-8 times in five sequences, and 0 times in one sequence, per stimulus duration condition. In total, there were about 30 appearances of different facial identities for each stimulus duration condition: each stimulus duration condition of the EEG frequency-tagging experiment was also tested in the behavioral experiment. The total testing session for the behavioral experiment lasted about 45 minutes.

### **EEG Acquisition**

EEG was acquired with a BioSemi ActiveTwo system (BioSemi B.V., Amsterdam, The Netherlands; for electrode coordinates: <http://www.biosemi.com/headcap.htm>; for electrode relabeling: Rossion et al., 2015, Fig. S1) with 128 Ag-AgCl Active-electrodes, sampled at a rate of 512 Hz. Further details are exactly as in Retter et al., 2020.

### **Analysis**

#### ***EEG Frequency-Tagging Experiment***

Data were analyzed with Letswave 5 (<https://www.letswave.org>) running on MATLAB R2013b (The MathWorks, USA).

#### *Preprocessing*

Drifts in offset during pauses of the recording were realigned to the pre-pause offset, and processed as in Retter et al. (2020). That is: bandpass filtered (Butterworth 4<sup>th</sup> order from 0.05 to 140 Hz); segmented from 2 s before to 80 s after stimulus presentation onset; corrected for muscular activity associated with eye blinks by the removal of single ICA component (for

1 participant blinking  $> 0.2$  times/s;  $M = 0.06$  blinks/s;  $SD = 0.072$  blinks/s); noisy channels were interpolated with 3-4 neighboring channels ( $M = 1.0$  channels;  $SD = 1.15$  channels; range = 0-4 channels); and referenced to the average of all 128 EEG channels.

### *Regions-of-interest*

The significance of FI responses at  $F/n$  was assessed over a ten-channel bilateral occipito-temporal (OT) region-of-interest (ROI), defined *a priori* (Rossion, Retter & Liu-Shuang, 2020). To probe the FI responses in further detail, the amplitude of the averaged right and left OT sub-regions were computed separately (right: channels P10; P8; PO8; PO10; PO12; left: channels P9; P7; PO7; PO9; PO11). The bilateral OT ROI was verified *post-hoc*: it captured 7 to 9 of the 10 channels producing the largest FI responses at the group-level across presentation conditions from 50 to 333 ms (20 to 3 Hz), and one channel at 25 and 33 ms (40 and 30 Hz). Across 50 to 333 ms (20 to 3 Hz), three right OT channels, P10, PO10, and PO12, were consistently defining two or three of the top three channel amplitudes. A region-free determinant of the FI response significance was also assessed over the average of all 128 EEG channels. To measure stimulus-presentation responses, a medial-occipital (MO) ROI was selected *a-priori* (channels O2; POI2; I2; Iz; OIz; Oz; POOz; O1; POI1; I1; Retter et al., 2020), and verified *post-hoc* to capture 9/10 of the maximal channels with a grand-average across stimulus durations; the average of all 128 EEG channels was also assessed.

### *Frequency Domain Analysis*

Data were processed for the frequency domain as in Retter et al. (2020). In brief: the 7-s stimulus durations steps were segmented in two ways: 1) from 100 ms before the first oddball identity onset, to capture FI responses and; 2) at oddball identity onset, to isolate the full stimulus-presentation response. Then, sequences were averaged within each frequency step condition and transformed into the frequency domain by means of a fast Fourier transform (FFT). Specific harmonic frequencies of the FI response ( $F/n = 1$  Hz) were summed up to 20 Hz; harmonics frequencies of the stimulus-presentation response ( $F$ ) were summed up to 60 Hz (frequency range: Retter et al., 2020; harmonic summation: Retter, Rossion & Schiltz, 2021). Six neighboring frequency bins determined the baseline (bl) “noise”. Significance of responses at the group and individual level were assessed with Z-Scores ( $Z = \text{signal} - \text{average}(\text{bl}) / \text{standard deviation}(\text{bl})$ ; significance threshold at 2.32,  $p < .01$ , 1-tailed testing signal  $>$  baseline noise) on the grand-averaged or individual average OT ROI channels for FI responses, and on the average MO ROI channels for stimulus-presentation responses.

Response amplitude was quantified by subtracting the average baseline noise (signal-average(bl)) from the summed-harmonic responses. Scalp topographies of these summed-

harmonic responses were visualized in terms of amplitude and normalized (according to the method of McCarthy & Wood, 1985) amplitude. A hemispheric lateralization index comparing the right (R) and left (L) OT sub-regions was calculated as follows:  $(R-L)/(R+L)*100$ . Note that average values for the right or left hemisphere that were below 0  $\mu$ V after noise-correction were corrected to 0  $\mu$ V before being input into the index. As in our previous study (Retter et al., 2020), to relate individuals' behavioral and neural responses, amplitude in the middle stimulus duration (50 to 125 ms; 8 to 20 Hz: see Results) was weighted by their amplitude at the longest duration (333 ms; 3 Hz) in order to normalize inter-individual EEG variability due to general factors (e.g., skull thickness, source orientation with respect to the scalp, etc.). Finally, statistical tests were performed on response amplitudes, independently for FI and stimulus-presentation responses, with one-way repeated measures analysis-of-variance (ANOVA) tests on the factor of *Condition*. Only adjacent stimulus duration steps were statistically compared in post-hoc analyses, with paired-sample, two-tailed, t-tests, with the application of a Benjamin-Hochberg correction for the ten allowed comparisons.

#### *Time Domain Analysis*

Segmented data were filtered more conservatively (Butterworth 4<sup>th</sup> order at 30 Hz); stimulus-presentation responses were removed with a FFT notch filter at the fundamental and harmonic frequencies up to 30 Hz; data were re-segmented as described above for measuring oddball FI responses; and averaged by stimulus duration condition (see Retter et al., 2020, for more details). Data were baseline offset-corrected with 100 ms preceding oddball stimulus onset, and significance was assessed over the right occipito-temporal sub-region with two-tailed t-tests relative to 0  $\mu$ V, with a threshold of  $p < .0001$ , across a minimum of 5 consecutive time bins (about 10 ms), that is, with strict criteria to reduce false-positives. Additionally, to more fully characterize the FI response, the data were averaged across all conditions producing significant frequency-domain individuation responses.

### ***Behavioral Experiment***

#### *Behavioral Face Individuation Analysis*

As above, and in Retter et al. (2020), responses to detecting occasional, non-periodic facial identity changes were considered correct when occurring between 0.15 and 2 s after identity change onset; responses outside this range were considered false positives. The total percent accuracy was corrected for false positives: the percent of false positives was subtracted from the percent of correct responses. Response time (RT) was calculated for correct trials, and inverse efficiency (IE = RT/accuracy) was calculated (Townsend & Ashby, 1983). Statistical tests were performed as for the neural responses: a one-way repeated measures ANOVA test

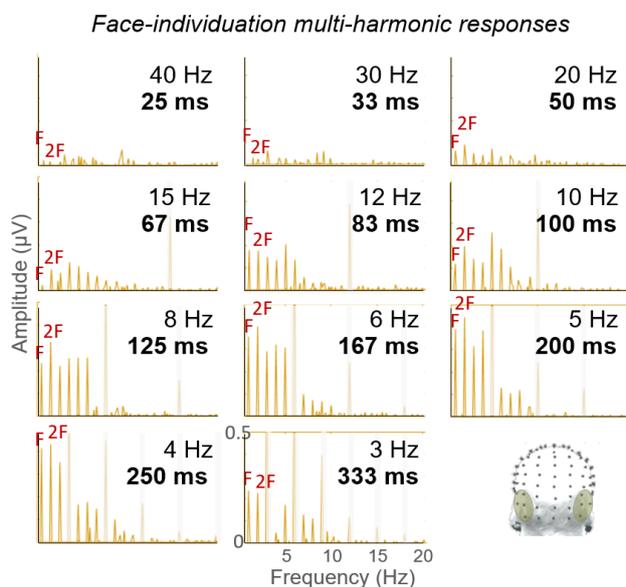
on the factor of *Condition*, and post-hoc t-tests on adjacent stimulus-presentation rates with the same criteria as given above. Finally, the relationship between behavioral FI accuracy and EEG amplitude were tested with Pearson correlations, two-tailed, both at the group and individual-participant levels, excluding outliers above 2.5 SD of the mean.

## Results

First, we investigated the minimal and optimal viewing times required for face individuation (FI), in terms of both implicit, EEG data and explicit, behavioral data. Next, we examined individual differences and the relationship between neural and behavioral data at the individual participant level.

### Minimal and Optimal Viewing Times for Face Individuation

In order to identify the *minimal* stimulus presentation time at which FI occurred, the shortest presentation duration producing significant responses was assessed both for neural and behavioral responses. Note that the neural response significance was assessed internally within each stimulus duration condition, on the sum of a range of unique harmonic frequency responses up to 20 Hz, with the tagged frequency bins vs. a range of local frequency noise (Retter, Rossion & Schiltz, 2021; see Methods for details; see **Fig. 3** for harmonic distributions of face-individuation frequency response amplitude).



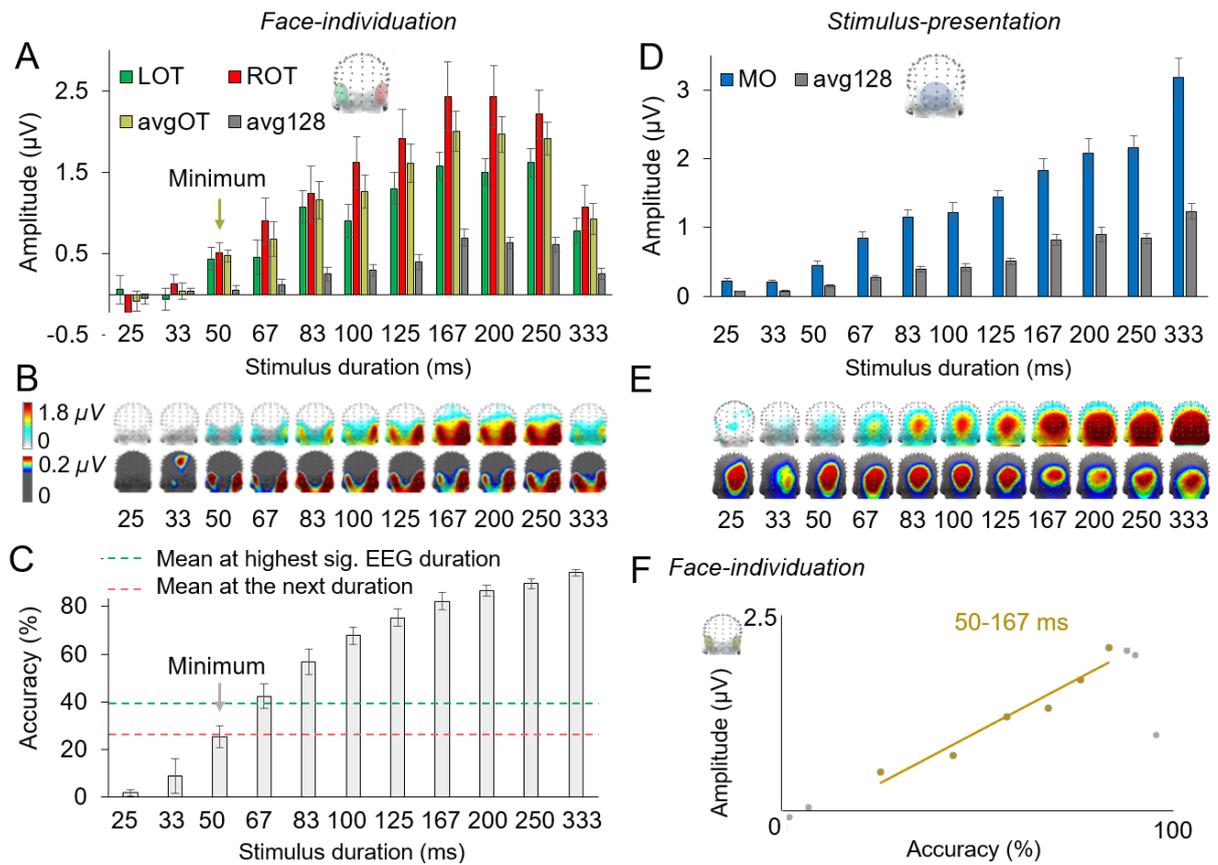
**Figure 3.** *Frequency-domain, baseline-subtracted amplitude spectra, showing harmonic frequency responses. Face individuation responses at 1 Hz and its harmonics, 2 Hz, 3 Hz, etc., up to 20 Hz, plotted from the bilateral occipito-temporal ROI. Harmonics coinciding with the stimulus-presentation rate are shadowed in light gray, and were excluded from face individuation response analyses.*

Neural response significance first emerged at 50 ms (20 Hz;  $p = .004$ ; **Table 1 A**) over the occipito-temporal ROI at the group level. At this rate, the amplitude was about 0.5  $\mu\text{V}$  (**Fig. 4 A**; **Table 2 A**; sub-region amplitudes and scalp topographies: **Fig. 4 A&B**). Strikingly,

behavioral group-level response significance also first emerged at 50 ms (20 Hz), at which duration the accuracy was about 25% (**Fig. 4 C; Table 2 B**). In contrast, stimulus-presentation responses were significantly present at all stimulus durations (**Fig. 4 D&E; Table 1 B**; see **Fig. S1** for harmonic stimulus-presentation frequency response amplitude distributions). For face individuation, there was a strong, positive linear correlation between accuracy and amplitude, particularly in a middle stimulus duration range, from 50 to 167 ms (20 to 6 Hz),  $r = .97$ ,  $p = .0014$  (**Fig. 4 F**).

Time (ms)	25	33	50	67	83	100	125	167	200	250	333
Frequency (Hz)	40	30	20	15	12	10	8	6	5	4	3
<b>A. Identity</b>											
OT	0.07	-0.68	<b>3.32</b>	<b>3.35</b>	<b>6.92</b>	<b>5.69</b>	<b>9.47</b>	<b>10.7</b>	<b>9.33</b>	<b>12.5</b>	<b>4.07</b>
Avg128	-0.16	0.15	0.40	0.70	1.78	1.83	<b>2.74</b>	<b>4.18</b>	<b>3.69</b>	<b>3.71</b>	1.14
<b>B. Stimulus</b>											
MO	<b>17.6</b>	<b>18.6</b>	<b>36.3</b>	<b>41.4</b>	<b>48.7</b>	<b>37.5</b>	<b>43.2</b>	<b>39.3</b>	<b>35.6</b>	<b>47.7</b>	<b>70.8</b>
Avg128	<b>20.6</b>	<b>24.3</b>	<b>39.1</b>	<b>35.7</b>	<b>28.9</b>	<b>16.3</b>	<b>23.7</b>	<b>39.1</b>	<b>29.3</b>	<b>26.4</b>	<b>30.0</b>

**Table 1.** Z-scores. Significant group-level responses at each stimulus presentation duration condition are shown in bold ( $Z > 2.32$ ,  $p < .01$ ); non-significant Z-scores are in plain type. Responses to the identity change (**A**) are shown for the bilateral occipito-temporal (OT) ROI and the average of all 128 channels (Avg128). Responses to stimulus presentation (**B**) are shown for the medial-occipital (MO) ROI and the average of all 128 channels.



**Figure 4.** Group-level EEG responses to face-identification (base-subtracted, summed harmonics of 1 Hz, up to 20 Hz) contrast to those to stimulus-presentation (base-subtracted, summed harmonics of  $F$  Hz, up to 60 Hz) across stimulus durations. **A)** Face individuation response EEG amplitude at the occipito-temporal ROI, as well as its left and right sub-regions, and the average of all 128 EEG channels. The minimum (shortest significant) stimulus duration for FI is indicated. Throughout, error bars indicate  $\pm 1$  SE of the mean. **B)** The corresponding FI scalp topographies. Upper row: Original amplitudes. Lower row: Normalized amplitudes. **C)** Behavioral face individuation response accuracy. The minimum (shortest significant) stimulus duration for FI is indicated, matching that of **Panel A**. The dotted green line indicates the mean accuracy across each individual's highest frequency producing a significant EEG response; the dotted red line indicates the mean accuracy at each individuals' preceding (non-significant) stimulus duration. **D)** Stimulus-presentation response EEG amplitude at the medial-occipital ROI and the average of all 128 EEG channels. **E)** Stimulus-presentation scalp topographies, lower row normalized; to the same scale as in **Panel B**. **F)** Identity-recognition EEG amplitude over the occipito-temporal ROI is linearly correlated with behavioral accuracy,  $r_4 = .97$ ,  $p = .0014$ , in a middle stimulus duration range from 50 to 167 ms (20 to 6 Hz; blue points). Each data point represents a stimulus-presentation duration condition (original data: **Panels A & C**). (Note that the correlation across all stimulus presentation

durations (25 to 333 ms; 40 to 3 Hz) was also significant,  $r_9 = .89$ ,  $p = .0002$ .) **Key)** MO: medial-occipital ROI; *avg128* = average of all 128 EEG channels.

In regards to the neural responses, a one-way repeated measures ANOVA revealed highly significant differences across stimulus durations over the occipito-temporal ROI:  $F_{10,150} = 15.0$ ,  $p < .001$ ,  $\eta_p^2 = 0.48$ . Post-hoc t-tests were performed only for the 10 adjacent stimulus duration conditions (see Methods). These tests revealed significant differences from 33 to 50 ms (30 to 20 Hz),  $t_{15} = 5.0$ ,  $p < .001$ ,  $d = 1.23$ , reflecting the difference between absent responses at 33 ms (30 Hz) and shorter durations, and significant responses at 50 ms (20 Hz) and longer durations. They also revealed significant differences where response amplitude increased from 67 to 83 ms, 100 to 125 ms, and 125 to 167 ms (15 to 12 Hz, 10 to 8 Hz, and 8 to 6 Hz; all  $p$ 's  $< .013$ ,  $d$ 's  $\geq .40$ ). No significant differences were present from 167 to 200 ms (6 to 5 Hz) or 200 to 250 ms (5 to 4 Hz;  $p$ 's  $> .7$ ,  $d$ 's  $\leq .06$ ). Finally, the response amplitude *decreased* from 250 to 33 ms (4 to 3 Hz),  $t_{15} = -5.76$ ,  $p < .001$ ,  $d = -1.24$ .

For behavioral responses, a one-way repeated measures ANOVA disclosed highly significant differences in accuracy across stimulus durations,  $F_{10,150} = 68.7$ ,  $p < .001$ ,  $\eta_p^2 = 0.81$ . Post-hoc t-tests on adjacent stimulus duration conditions showed significant accuracy differences on the pairs from 25 to 33 ms (40 to 30 Hz) and 83 to 100 ms (12 to 10 Hz; all  $p$ 's  $\leq .004$ ,  $d$ 's  $> .44$ ). At durations longer than 100 ms (10 Hz), there were no significant accuracy differences across adjacent durations, although two further comparisons neared significance, 125 to 167 ms (8 to 6 Hz),  $p = .033$ , critical value = .023,  $d = .44$ , and 250 to 333 ms (4 to 3 Hz),  $p = .035$ , critical value = .027,  $d = .65$ ; all other  $p$ 's  $> .067$ ,  $d$ 's  $\leq .47$ . Note that across 100 to 333 ms (10 to 3 Hz), participants' response times remained similar (ranging between 500 to 517 ms, with the minimum of 500 ms occurring at 125 ms (8 Hz; **Table 2 C**). However, the accuracy across this range increased by nearly 40%, albeit in relatively small increments.

Time (ms)	25	33	50	67	83	100	125	167	200	250	333
Frequency (Hz)	40	30	20	15	12	10	8	6	5	4	3
<b>A. Amplitude</b>	-0.08 (0.12)	0.04 (0.10)	0.48 (0.07)	0.68 (0.21)	1.16 (0.23)	1.26 (0.20)	1.61 (0.24)	2.01 (0.25)	1.97 (0.22)	1.92 (0.21)	0.93 (0.19)
<b>B. Accuracy</b>	1.84 (1.45)	6.30 (4.92)	25.5 (4.61)	42.4 (5.11)	56.8 (5.45)	67.8 (3.68)	75.3 (3.57)	82.1 (3.63)	86.7 (2.18)	89.6 (2.27)	94.2 (1.29)
<b>C. RT</b>	670 (95.0)	510 (45.2)	570 (23.7)	557 (20.9)	540 (22.2)	517 (15.5)	500 (13.0)	503 (11.5)	514 (10.6)	514 (13.7)	514 (11.7)

**Table 2.** *Face individuation performance: A) EEG amplitude over the bilateral occipito-temporal (OT) ROI; B) percent accuracy; and C) correct response time (RT). Standard error ( $\pm 1$ ) across participants is indicated in parentheses.*

To summarize, the minimal stimulus presentation duration for both neural and behavioral FI responses was 50 ms (20 Hz). The optimal duration for neural responses was at 167 ms (6 Hz), with practically equivalent amplitudes present from 167 to 250 ms (6 to 4 Hz). At 167 ms, accuracy was high (82%), and although it continued to increase until the optimal, longest stimulus duration of 333 ms, there were no significant differences in adjacent rate steps between 100 to 333 ms (10 to 3 Hz), with similar response times. Finally, group-level accuracy and amplitude were strongly correlated across stimulus durations.

## **Individual Differences in Face Individuation**

### ***Stimulus Presentation Rate Diagnosticity***

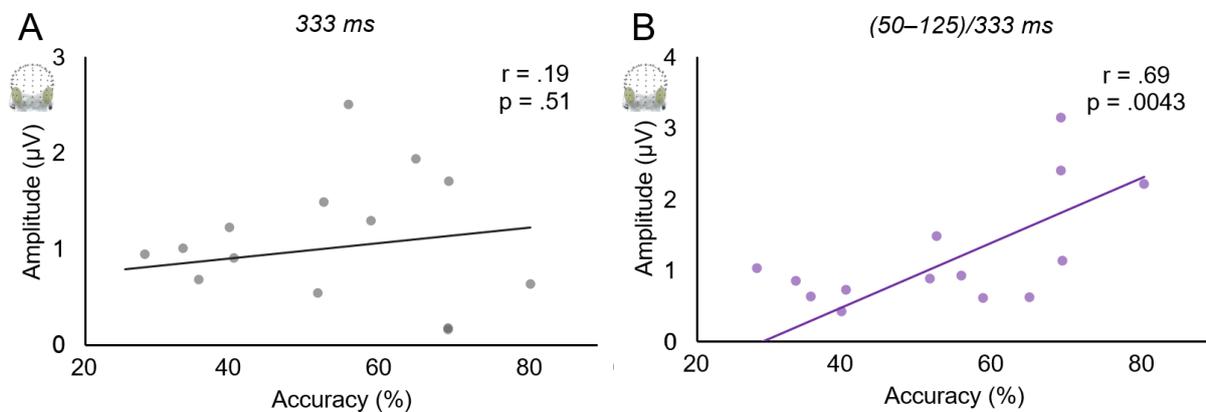
There were substantial inter-individual differences in the minimum stimulus presentation duration required for participants to individuate unfamiliar faces (please see the last section of the results). However, at the shortest stimulus presentation durations (25 and 33 ms; 40 and 30 Hz), there was not a large amount of inter-individual variability: group-level amplitude was not above zero at these rates, and amplitude was very low for nearly all participants (**Fig. 4 A; Fig. S2 A**). Behaviorally, only at most a few identity changes were detected by any participant at these shortest durations (**Fig. 4C; Fig. S2 B**).

At the other extreme, there were also only small inter-individual differences at the longest durations, at which performance at FI was generally high (although not at ceiling). From 125 to 333 ms (8 to 3 Hz), group-level accuracy was consistently above 80% (**Table 2 B**; average SE = 2.4%; range across participants progressively increasing from 16 to 47%). The largest separation across individual participants' behavioral performance was thus in a middle stimulus duration range, from 50 to 125 ms (20 to 8 Hz), wherein the group-level accuracy ranged from about 75 to 25% (average SE = 4.5%; range across individual participants consistently above 62%). Thus, to examine inter-individual differences, individuals' accuracy was averaged across conditions from this middle range (50 to 125 ms; 8 to 20 Hz). This middle range was also used to examine EEG responses in the frequency domain, where amplitude was significant but below the optimal value (**Fig. 4 A; Fig. S2 A; Fig. S3** for individuals' scalp topographies in this range).

### ***Individuals' Behavioral Accuracy Correlates with Neural Response Amplitude in a Weighted Middle Stimulus Duration Range***

The diagnostic middle stimulus duration range (50 to 125 ms; 8 to 20 Hz) was used to explore the relationship between individuals' behavioral performance and EEG amplitude from the frequency domain. Additionally, the amplitude at the longest stimulus duration, 333 ms (3 Hz), was used as a baseline amplitude measure for each participant to remove inter-individual variability due to general factors (see Methods). Specifically, we tested the hypothesis that individuals' 3-Hz weighted amplitude in the middle stimulus duration range correlates with their behavioral performance in the middle stimulus duration range.

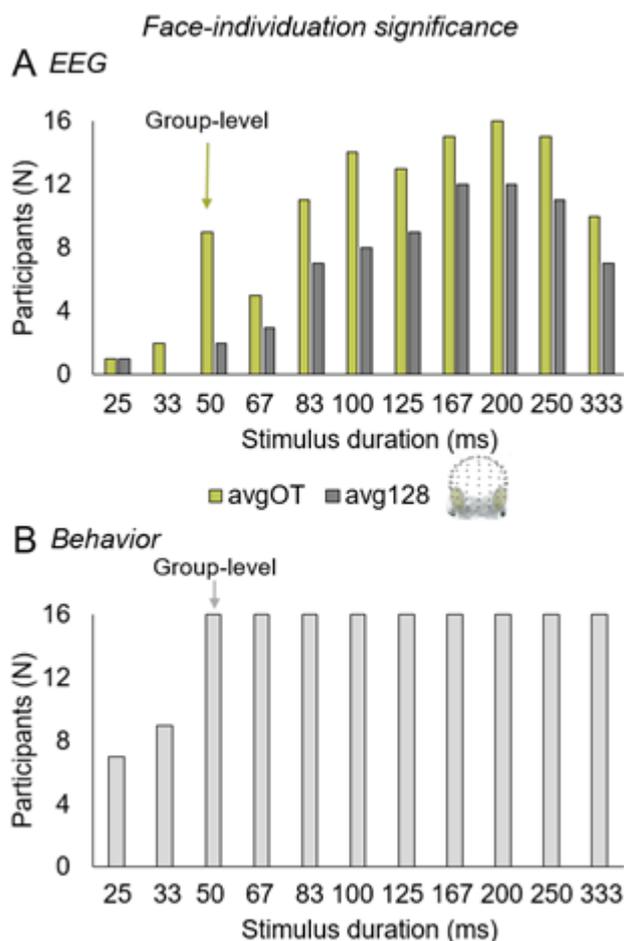
As expected, there was no correlation between individuals' behavioral performance and EEG amplitudes at 333 ms (3 Hz), slope = 0.44  $\mu\text{V}/\%$ ,  $r_{13} = .19$ ,  $p = .51$  (**Fig. 5 A**). However, there was a significant, positive linear correlation between individuals' behavioral performance and EEG amplitudes in the weighted middle duration range (the average of 50 to 125 ms divided by 333 ms), slope = 1.8  $\mu\text{V}/\%$ ,  $r_{13} = .69$ ,  $p = .0043$  (**Fig. 5 B**). Note that similar results were obtained when comparing EEG amplitude with inverse efficiency, a combined measure of individuals' response time and accuracy (**Fig. S4**).



**Figure 5.** Correlations between individual participant face individuation accuracy (behavioral experiment: 50 to 125 ms; 20 to 8 Hz) and face individuation frequency-domain amplitude (EEG frequency-tagging experiment: presentation rates as indicated; base-subtracted, summed harmonics of 1 Hz, over the OT ROI). **A)** The non-significant correlation of individuals' accuracy with amplitude at the longest stimulus duration, 333 ms (3 Hz). **B)** The significant correlation of individuals' accuracy with amplitude in the middle stimulus duration range, 50 to 125 ms (20 to 8 Hz), weighted by the amplitude at 333 ms (3 Hz).

### ***Neural Response Latency Predicts Individuals' Behavioral Accuracy***

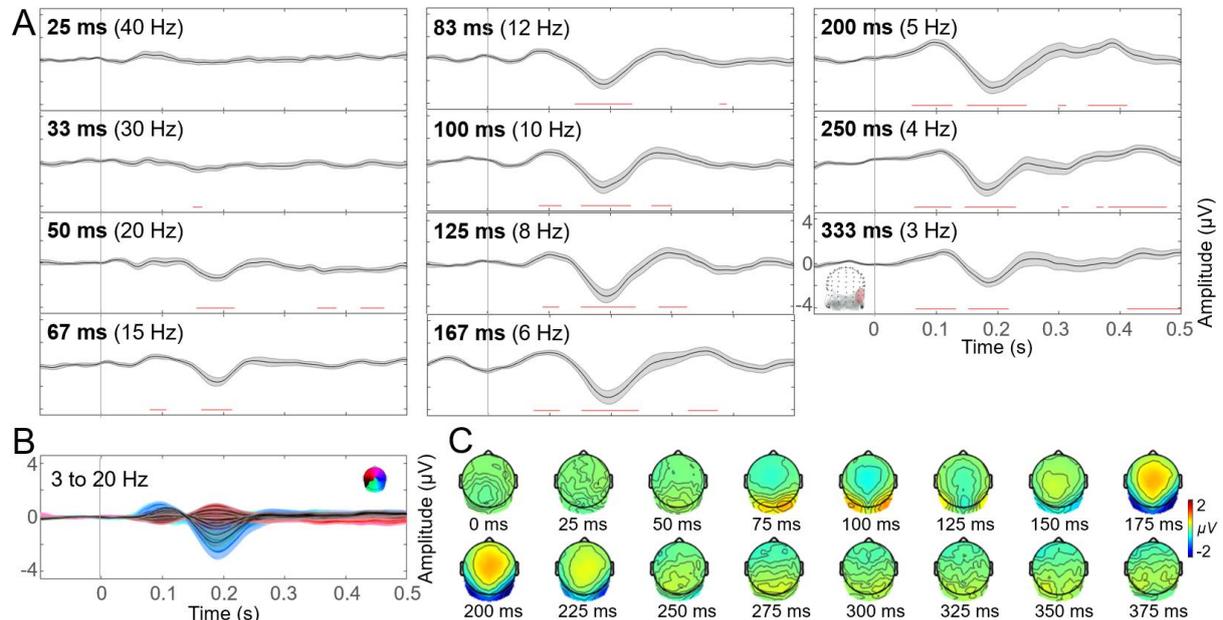
Individual differences were present in terms of the earliest significant FI EEG responses, which occurred across a range of 25 to 100 ms (40 to 10 Hz) over the occipito-temporal ROI in the frequency domain (**Fig. 6 A**). At the high end, only one participant had a significant EEG response at the shortest stimulus presentation duration (25 ms; 40 Hz). Two participants had significant responses at the next duration, 33 ms (30 Hz), and 9 at the following duration, 50 ms (20 Hz). In terms of behavioral response accuracy, the earliest significant individual responses occurred across a range of 25 to 50 ms (40 to 20 Hz; **Fig. 6 B**). Seven participants had significant behavioral FI responses at the shortest stimulus presentation duration (25 ms; 40 Hz), nine at the next duration (33 ms; 30 Hz), and all participants at the following duration and below (50 ms; 20 Hz).



**Figure 6.** Significance of individual-participant face individuation responses across presentation rates, for both EEG frequency-tagging and behavioral experiments. **A)** The number of participants (out of 16) with a significant occipito-temporal EEG response at each rate. Group-level significance first appeared at 50 ms (20 Hz). **B)** The number of these participants with more correct than false alarm behavioral responses at each rate. **Key)** avgOT: bilateral occipito-temporal ROI; avg128 = average of all 128 EEG channels.

To assess whether the latency of individuals' neural FI responses predicts individual differences in the viewing time of FI behaviorally, we related individuals' accuracy (in the middle stimulus duration range, 50 to 125 ms; 20 to 8 Hz) to the peak latency of the first and second deflections, which have been described in previous studies (Rossion, Retter & Liu-

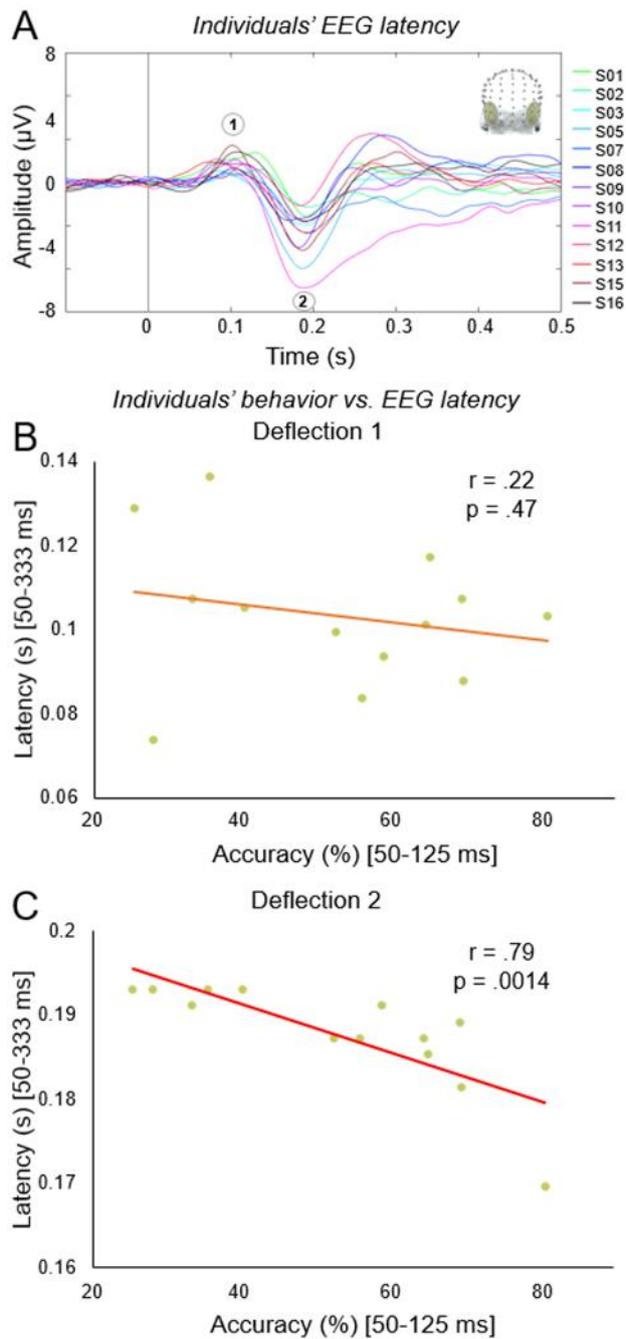
Shuang, 2020). Note that given the high consistency of the temporal dynamics of FI responses across rates (**Fig. 7 A**; see also **Fig. S5** for non-*F*-filtered data), stimulus duration rates were combined from 50 to 333 ms (20 to 3 Hz), i.e., all conditions producing significant FI responses in the frequency domain, in order to increase the signal-to-noise ratio by increasing the number of identity change events (resulting in more than 1,000 events per participant; see Methods). These first two deflections: 1) a first, positive deflection, first reaching significance at 74 ms, and peaking at approximately 100 ms at the group level; and 2) a second, negative deflection, first reaching significance at 152 ms, and peaking at approximately 185 ms at the group level (**Fig. 7 B&C**), were readily identifiable here in thirteen out of sixteen participants (**Fig. 8 A**; see also **Fig. S6 A**). To capture the most sensitive individual responses in the time domain, measurements were made from either the right or left occipito-temporal sub-region, corresponding to the lateralization index of each participant from the frequency-domain analyses, from 50 to 333 ms (20 to 3 Hz;  $M = 0.13$ , i.e., higher activation over the right hemisphere;  $SE = 0.071$ ; range =  $-.30$  to  $.70$ ; see again **Fig. S3** for individual scalp topographies); and note that similar results were produced when using the bilateral occipito-temporal ROI (**Fig. S6 B&C**).



**Figure 7.** Temporal dynamics of the face individuation response, showing similar deflections over time across stimulus duration steps. **A)** Time-domain identity recognition responses over the right occipito-temporal sub-region for each stimulus duration condition. Significant deflections are indicated with a red line below. There were three prominent deflections: 1) a positive deflection, peaking at approximately 100 ms post-identity change onset, in conditions

at which it reached significance (highlighted in orange;  $SD = 6.9$  ms; significance window: about 75 to 120 ms, across these conditions); 2) a second, negative deflection, peaking at approximately 185 ms (highlighted in yellow;  $SD = 10.6$  ms; significance window: about 150 to 225 ms); and 3) a third, positive deflection, peaking at approximately 295 ms (highlighted in green;  $SD = 13.0$  ms; significance window: about 280 to 315 ms). Note that these waveforms have been notch-filtered in the frequency domain to remove the stimulus-presentation responses. **B)** The identity-recognition response averaged across frequency-domain significant conditions, from 50 to 333 ms (20 to 3 Hz). For this data, the first, positive deflection of the right occipito-temporal sub-region again peaked at approximately 100 ms, with an amplitude of  $0.85 \mu\text{V}$ . The negative peak followed at approximately 190 ms, with an amplitude of  $-2.3 \mu\text{V}$ , followed by the third, positive peak at about 290 ms, having an amplitude of only  $0.18 \mu\text{V}$ . **C)** Two-dimensional scalp topographies of the data shown in Panel B, sampled every 25 ms from 0 to 375 ms post-identity change onset.

For the first, positive deflection, the mean post-identity change peak latency across individuals was 102 ms ( $SE = 4.7$  ms), with a range of 47 ms. There was an insignificant, weak, negative correlation between participants' behavioral accuracy and the latency of this first deflection, slope =  $-0.29$  ms/%,  $r_{11} = .30$ ,  $p = .32$  (**Fig. 8 B**). For the second, negative deflection, the mean latency across individuals was 190 ms ( $SE = 1.3$  ms), with a range of 14 ms. In this case, however, there was a significant, negative correlation between individuals' accuracy and EEG latency of the second deflection, slope =  $-0.21$  ms/%,  $r_{11} = .85$ ,  $p = .00027$  (**Fig. 8 C**). That is, participants who had significantly faster neural responses at the time of the second, negative deflection also performed better at individuating faces across stimulus presentation durations from 50 to 125 ms (20 to 8 Hz). As a reminder, electrophysiological and behavioral measures were taken in separate experiments.



**Figure 8.** The relationship between individuals' EEG peak latencies (across stimulus durations producing significant face individuation responses, from 50 to 333 ms; 20 to 3 Hz); and behavioral performance (in the middle stimulus duration range, from 50 to 125 ms; 20 to 8 Hz). Data were plotted at either the right or left occipito-temporal sub-region, corresponding to the lateralization index of each participant from the frequency-domain analyses. Note that three participants were removed from this analysis for not showing clear response peaks at the time of these deflections (**Fig. S6 A**). **A**) Individual EEG peak latencies were measured in the time domain, wherein periodic facial identity changes were presented every 1 s (1 Hz; in this Panel, 0 s = identity change onset). **B**) There is a non-significant, negative correlation between individuals' behavioral accuracy and peak latency of the first, positive deflection. **C**) There is a significant, negative correlation between individuals' behavioral accuracy and peak latency of the second, negative deflection.

## Discussion

### The Minimal Stimulus Viewing Time for Face Individuation

As reviewed in the Introduction, behavioral and neural studies carried out independently have struggled to provide consistent answers to questions concerning cognitive functioning, such as: how much viewing time does it take to individuate a complex visual stimulus, such as a face? Here, we measured neural activity and behavior with the same participants, using parallel

experiments comprised of varying facial identity viewing times. A group-level neural face individuation (FI) response emerged at 50 ms stimulus viewing time (20 Hz; **Fig. 4 A**). The same viewing time, 50 ms (20 Hz), was also found for the emergence of behavioral FI (**Fig. 4 C**).

This minimal duration is shorter than the 120 ms (8.33 Hz) duration reported for neural FI by Alonso-Prieto et al. (2013), when contrasting EEG responses at the stimulus presentation rate to periodic sequences of different (i.e., a temporal crowd) vs. same (i.e., a single repeating) facial identities. This may be accounted for by differences in experimental design. Here, FI responses across stimulus durations are both measured within each stimulation sequence, since the FI responses occur at a separate frequency than the stimulus presentation responses. Importantly, the FI responses are non-overlapping here, as stimuli are 1 s apart, regardless of the stimulation rate. In particular, overlapping neural responses to different faces may have caused the lack of significant effects above 8.33 Hz in that previous study, in which FI responses were measured at the stimulation rates (but not in Gentile & Rossion, 2014 with fMRI; see Keyers & Perrett, 2002; Keyers et al., 2005; Henson, 2016; Retter et al., 2016; Retter et al., 2018; Retter et al., 2020). The minimal duration of 50 ms is longer than that of several previous behavioral studies, reporting FI at durations below 20 ms when not applying backward masking (Bachmann, 1991; Gur, 2018), or when using only a few, highly familiar stimuli (Rolls et al., 1994). On the other hand, the present behavioral results are in near agreement with Or & Wilson (2010), reporting a minimum of about 60 ms (to achieve 75% accuracy at a two-alternative force-choice task, with morphed, synthetic stimuli devoid of surface cues) and in full agreement with the 50 ms duration reported by Tankanen et al. (2007), using natural images of 6 familiarized male faces.

At this 50 ms duration, only about 1 in 4 identity changes were correctly reported behaviorally at the group level. Thus, this 50 ms minimum is not an absolute limit: at shorter stimulus durations, faces may still be individuated, only less often; at longer stimulus durations, faces may be individuated more reliably. Further, this value is approximate here, in the sense that no stimulus durations were tested in between 33 and 50 ms (30 and 20 Hz).

### **The Optimal Stimulus Viewing Time for Face Individuation**

After emerging, neural FI responses at  $F/n$  at the group-level continued to increase over a middle stimulus duration range, from 50 to 125 ms (20 to 8 Hz), first reaching their maximum at 167 ms (6 Hz; **Fig. 4 A**; **Table 2 A**). This 167 ms duration (6 Hz) value is defined as the *optimal* stimulus duration, and is in line with the previous maximal, differential, neural FI amplitudes at 6 Hz reported by Alonso-Prieto et al. (2013) and Gentile & Rossion (2014), with

EEG and fMRI, respectively, for sequences of different *vs.* same facial identities (see also Rossion, 2014). This also is roughly in line with the durations for MEG responses reported by Tanskanen et al. (2007) for face detection, that increased from durations of 100 to 200 ms, although durations beyond 200 ms were not tested.

At durations shorter than 167 ms (6 Hz), progressively lower FI response amplitudes may be accounted for by missed FI, i.e., not through gradually reduced neural responses, but through the proportion of absent or present all-or-none neural responses (Retter et al., 2020). Such an effect could not be directly tested here, but is supported by the strong, positive linear correlation between amplitude and accuracy, from 24 to 6 Hz (**Fig. 4 F**; see also Kovacs, Vogels & Orban, 1995; Vanni et al., 1996; Grill-Spector et al., 2000; Bacon-Mace et al., 2005; Retter et al., 2020). Further in line with this interpretation, responses were qualitatively similar across variable stimulus durations producing differing amplitudes, rather than being partially degraded with less viewing time (scalp topographies: **Fig. S3**; time-domain dynamics: **Fig. 7**). Whether or not each oddball face was individuated at each presentation in this paradigm is likely influenced by its difference from the contrasted base facial identity, which could be defined in terms of multiple physical cues, as well as the experience and ability of the individual observer.

Here, at durations longer than 167 ms (6 Hz), neural response amplitude plateaued from 200 to 250 ms duration (5 to 4 Hz), beyond which there was a *decreased* neural FI amplitude at the longest 333 ms stimulus duration (3 Hz; **Fig. 4 A**; **Table 2 A**). The decreased, differential FI amplitudes may be caused by less decreased responses to the base face here, by means of: 1) reduced neural interference of base stimulus responses at 333 ms, since this inter-stimulus interval is longer than that of the bulk of FI responses observed here in the time domain (with significant responses of the third deflection offsetting at about 240 ms from the onset of significance of the first deflection; **Fig. 7**; see also Retter et al., 2020, for a slightly decreased response at 3 Hz); and 2) reduced repetition suppression to the base face, because there are only two presentations (one repetition) of the base face for each presentation of the oddball face, i.e., more frequency identity changes than identity repetitions, and a duration long enough for multiple gaze fixations (see Grill-Spector & Malach, 2001; Rossion, Retter & Liu-Shuang, 2020). Note that at low stimulus presentation frequencies (1 and 2 Hz), no difference between different *vs.* same facial identity responses were reported by Alonso-Prieto et al. (2013), while such effects have been reported at frequencies between about 3.5 and 8 Hz (Rossion & Boremanse, 2011; Alonso-Prieto et al., 2013; Nemrodov, Jacques & Rossion, 2015).

Behaviorally, the FI accuracy was highest at the longest stimulus duration, 333 ms (3 Hz; **Fig. 4 C**; **Table 2 B**), with stable response times from 100 to 333 ms (10 to 3 Hz; RTs ranging from a minimum of 500 ms at 125 ms duration (8 Hz) to a maximum of 517 ms at 100 ms duration (10 Hz; **Table 2 C**). This result contradicts that of previous behavioral studies, reporting optimal FI responses at much shorter stimulus durations (typically from about 35 to 110 ms: Rolls et al., 1994; Lehky, 2000; Tanskanen et al. 2007; Or & Wilson, 2010; and 4 ms in Bachmann, 1991). However, as addressed above, these previous studies did not apply backward masking (Bachmann, 1991; see also Gur, 2018), used only few, familiar(ized) face stimulus exemplars (Rolls et al., 1994; Tanskanen et al. 2007), or presented the exact same test and target images with discrete morphed changes (Lehky, 2000), such that the individuation task demanded may have been too easy, reaching ceiling performance at relatively short durations. Here, oddball stimuli (about 30 per stimulus duration) were randomly chosen from among 99 different facial identities (excluding the base identity in each sequence), and the image size and luminance varied at each presentation. Most importantly, contrary to previous studies, the number of oddball identity targets varied across sequences, with some sequences containing no oddballs. Hence, behavioral performance was never at ceiling in our task, only nearing it at the longest stimulus duration.

### **The Relationship Between Individuals' Neural and Behavioral Responses**

Inter-individual differences in stimulus processing time for FI were most prominent when the task was reasonably challenging, in the middle stimulus duration range, of 50 to 125 ms (20 to 8 Hz; **Fig. S2**). The FI amplitude in this range, weighted by the FI amplitude at 333 ms (3 Hz), significantly correlated with individuals' behavioral performance, measured independently (**Fig. 5 B**; see also Retter et al., 2020). Note that at long stimulus presentation durations, individual differences in behavior were weaker (**Fig. S2 B&C**). However, in most previous studies, explicit behavioral measures of FI typically rely on a relatively long, sometimes even unlimited, viewing times of unfamiliar face pictures (e.g., Duchaine & Nakayama, 2006). One reason for this is that time pressure in explicit tasks with unfamiliar FI can deteriorate behavioral performance (Bindemann et al., 2016; Fysh & Bindemann, 2017). Yet, although providing more time to individuate faces can lead to improvements of behavioral responses, this may not be ideal, as it may also lead to unnatural (e.g., analytical, feature-based) processes, making the measure less specific and diagnostic. In addition, while measures of behavioral response speed (time taken) are relevant for measuring individuals' ability (Rossion & Michel 2018; Dzhelyova et al., 2020), response time is not a good proxy of processing time, as it may be affected by many post-perceptual factors.

### Stimulus Viewing Time vs. Neural Response Latency

While we primarily investigated FI in terms of the minimal (and optimal) stimulus duration required, we were also able to examine EEG FI processing in the time domain (**Fig. 7**, having selectively filtered out stimulus-presentation responses (**Fig. S5**); see also Rossion, Retter & Liu-Shuang, 2020). Previous EEG studies taking this approach have produced conflicting results: it remains debated whether FI is captured at the (peak of the) occipito-temporal face-sensitive N170 component (Heisz et al., 2006; Jacques & Rossion, 2007; Nemrodov et al., 2019), on post-200 ms components such as the N250 (Schweinberger & Neumann, 2016), or even possibly at earlier latencies (e.g., Seeck et al., 1997; Nemrodov et al., 2016; Dobs et al., 2019).

Here, group-level, EEG FI responses were first observed at a positive deflection attaining significance at approximately 75 ms and peaking at approximately 100 ms post-identity-change onset (**Fig. 7**). While this could be taken as evidence for an early FI response onset, we argue *against* this account, in light of both previous and current indications. Importantly, in previous studies, several lines of evidence point to this deflection reflecting low-level stimulus changes, rather than high-level FI: it is disproportionately affected by stimulus size-change ranges and orientation, and it shows no face inversion effect, i.e., occurring with equal latency and magnitude for upright or inverted face images (Dzhelyova & Rossion, 2014; Rossion, Retter & Liu-Shuang, 2020). Additionally, in standard ERP paradigms, the P1(00) deflection occurring at about this latency is absent when low-level image changes are minimized (Jacques & Rossion, 2006; Zheng et al., 2011; compare also Nemrodov et al., 2019, to Nemrodov et al., 2016), it is not selective to faces beyond low-level cues contained in amplitude spectrum (Rossion & Caharel, 2011; Ganis et al., 2012) and its amplitude and latency do not correlate with group-level behavioral performance at FI across orientations, unlike the N170 (Jacques & Rossion, 2007).

In the present paradigm, this first, positive deflection may also be considered in light of a potential release from adaptation to the base face, as addressed above. This adaptation may occur both at a high-level, to facial identity, but also to low-level image attributes. This is a likely source of modest low-level contributions to the FI response measured here, despite the large stimulus set (100 different, unfamiliar facial identities), and continuous forward- and backward stimulus masking. Empirically, the peak latency of the first, positive deflection, was *not* significantly correlated with individuals' behavioral FI accuracy at middle stimulus durations (**Fig. 8 B**). Additionally, the scalp topography of this first, positive deflection, before

its peak, i.e., at about 50-75 ms, is centered over medial occipital channels, supporting a low-level interpretation (**Fig. 7 C**; Liu-Shuang, Retter & Rossion, 2020).

Thus, we suggest that FI responses that go beyond physical stimulus differences first occur only at the time window of the second, negative deflection, attaining significance at approximately 150 ms, and peaking at approximately 185 ms, here. This deflection first emerged over occipito-temporal channels (**Fig. 7 C**), and, importantly, its peak latency *did* significantly correlate with, i.e. predict, individuals' behavioral FI accuracy in a middle stimulus duration range (addressed in the following section; **Fig. 8 C**). This onset of selective FI responses is in line with some earlier studies, as addressed above (Heisz et al., 2006; Jacques & Rossion, 2007; Caharel et al., 2009; see Rossion & Jacques, 2011 for review), and with a view in which the early signatures of FI are present within the same time window as generic face (vs. object) categorization, i.e., by the onset at about 130 ms of the N170 component (Jacques & Rossion, 2006).

### **Neural Response Peak Latency (~185 ms) Predicts Individuals' Behavioral Face Individuation Ability**

At the individual level, the peak latency of the second, negative deflection, peaking at 185 ms on average at the group level over the right occipito-temporal cortex (and ranging from 183 to 197 ms at the individual participant level), significantly, negatively correlated with participants' behavioral FI responses in a middle stimulus duration range (accuracy ranging from 81% to 28%; **Fig. 8 A&C**): the earlier the neural response, the better the behavioral performance under time constraints. Here, accuracy at short viewing times was a more sensitive indicator of FI processing time than response time (RT), which varied relatively little across participants (**Fig. S2 C**; compare also **Fig. 5** with **Fig. S4**, showing similar results for accuracy and inverse efficiency, i.e., RT/accuracy), and is influenced by many general factors (e.g., decisional time, motor speed, etc.). This relationship may be reflective of high-level FI responses, since: 1) a correlation was *not* significantly present for the first, positive deflection peak latency; and 2) there was also *not* a consistent relationship observed between individuals' stimulus presentation response phase at *F* and behavioral performance (data not shown).

In general, the timing information of EEG responses has traditionally been related to group-level behavioral responses through either early-stage visual processing (e.g., stimulus contrast and spatial frequency: Strasburger, Scheidler & Rentschler, 1988), or attentional modulation (e.g., van den Berg et al., 2016). Previous studies investigating facial processing have reported ERP latency differences, but more often a lack of differences, across individuals.

Herzmann et al. (2009) reported that individuals' behavioral accuracy at a battery of face processing tasks was moderately, negatively related to the latency of the N170; however, an equal correlation was present for measures of object processing, such that this effect was attributed to general cognitive factors. Das, Giesbrecht & Eckstein (2010) reported no relationship between individuals' behavioral performance at detecting faces *vs.* cars in noisy images and N170 peak latency over a right occipito-temporal channel. There was no relationship observed between individuals' viewpoint discrimination performance improvement at a trained view and N170 latency reduction for that view (Su et al., 2012). At a group level, while there was no delay reported for the N170 peak latency to faces (*vs.* cars) for people with poor FI ability relative to age-matched controls (Towler et al., 2012; but for contradictory results with MEG: Lueschow et al., 2015), in a different experimental design there was delay of 35 ms of the N250 ERP to target faces (but not own faces) for the poor face recognition group, over a right occipito-temporal channel (Parketny, Towler & Eimer, 2015; for review, Towler et al., 2017).

However, in none of these studies, or other studies beyond face processing to our knowledge, has a relationship been observed between the latency of neural responses (directly related to a given function) and inter-individual variability (in a homogenous population) in processing time at this function, as found here. Importantly, the behavioral responses measured here are accuracy rates reflecting the ability of individuals to individuate faces more often at short viewing times (here, measured from 50 to 125 ms of face stimulus duration). Thus, the straightforward interpretation of our results is that *individuals with shorter neural response latency are also able to behaviorally individuate faces with shorter viewing times*. While it is possible that individuals who have faster FI processing are also better at FI more generally, e.g., having more extensive cortical populations involved in FI in the ventral occipito-temporal cortex, such a relationship remains unknown.

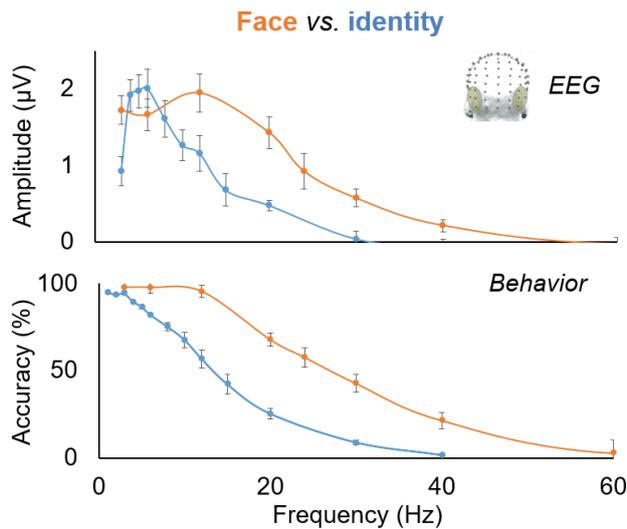
### **High-Level and Generalizable Face Individuation Responses?**

To what extent do the measures of FI reported here reflect *high-level*, face-specific processes? As mentioned previously, a number of previous studies using this paradigm point specifically to high-level responses (see Rossion, Retter & Liu-Shuang, 2020, for review). Additionally, there are several aspects of the present experimental design that may further limit contributions from low-level stimulus changes. Here, at every stimulus presentation there were substantial changes not only in size (from 80 to 120% of the original, sampled in 5% steps), but in luminance, which was varied from -10% to +10% of the original, sampled in 2.5% steps. Accordingly, identity changes could not be reliably detected from stimulus brightness and

contour position, resulting in below-ceiling individuation responses, even for stimulus viewing times of up to 333 ms (**Fig. 4 C**; **Table 2 B**).

In these experiments there were 100 facial identity stimuli (as in previous oddball FI studies, derived from color photographs of cropped, full-front, color, female faces, minimally standardized and cropped to the outer contour of the face; Liu-Shuang et al., 2014; Xu et al., 2017; Yan, Liu-Shuang & Rossion, 2019), leading to increased variability of oddball identities, i.e., reduced diagnosticity of specific low-level oddball *vs.* base differences. Empirical evidence of high-level responses here derives from right lateralized, occipito-temporal EEG responses, even at the shortest significant duration of 50 ms (20 Hz; contrasting with the medial-occipital responses general to visual stimulation: compare **Fig. 4 B** with **Fig. 4 E**), and the absence of both behavioral and neural response significance at the very short stimulus durations of 25 and 33 ms (40 and 30 Hz), when neural responses to stimulus-presentation are still recorded. However, the presence of a very early FI response in the time domain, i.e., with an onset much before 100 ms, suggests that there are some contributions of low-level effects on the FI response, perhaps resulting from low-level adaptation to the base face stimulus. Reduction or elimination of these effects with different stimulus sets or paradigms, or through the subtraction of responses to inverted faces (Rossion, Retter & Liu-Shuang, 2020; Jacques et al., 2020) could be examined in future studies.

Yet, for a point of reference, the 50 ms minimum stimulus duration for FI reported here is longer than that which has typically been reported for generic face categorization (i.e., about 15-35 ms), measured behaviorally and neurally by contrasts of faces *vs.* objects or phase-scrambled images (e.g., Or & Wilson, 2010; Mohsenzadeh et al., 2018). In a frequency-sweep design as applied here, categorization of a stimulus as a face among a rapid periodic train of non-face objects, excluding an effect based on amplitude spectrum differences between these stimuli (Rossion et al., 2015; Gao et al., 2018), a few (i.e., 3%) correct behavioral face categorization responses can be made even at 17 ms (60 Hz) duration, with a much larger rate of correct behavioral detections and significant neural measures observed at 33 ms duration (30 Hz; Retter et al., 2020). That is, FI responses here emerged with a delay of approximately 20 ms (50 *vs.* 33 ms, i.e., 20 *vs.* 30 Hz) relative to the minimum, but at least 80 ms (167 *vs.* 83 ms, i.e., 6 *vs.* 12 Hz) relative to the optimal, i.e., peak amplitude, generic face categorization responses (Retter et al., 2020; **Fig. 9**).



**Figure 9.** A comparison of the present group-level face individuation results (blue) with that of the frequency sweep of Retter et al. (2020; orange) for generic face (vs. object) categorization. In the previous study, non-face objects appeared as base stimuli, with faces appearing at the  $F/n = 1$  Hz rate.

This comparison should be made with a grain of salt: while the frequency-sweep design was similar, with target stimuli appearing at  $F/n = 1$  Hz (1 s), there were considerable differences in the experimental paradigm: the generic face categorization study used natural images with sometimes off-centered faces and non-face objects, and the “base” object images changed at every stimulus presentation and were highly variable (e.g., consisting of plants, man-made objects, animals, etc.; see Rossion et al., 2015). Still, a longer delay between the minimal and optimal durations required for these processes is in line with the view of a slower accumulation of sensory evidence, less resistant to degradation, for FI than for generic face categorization (e.g., Sergent, 1986; Crouzet et al., 2010; Or & Wilson, 2010; Amihai et al., 2011; Besson et al., 2017). Note that this view is not necessarily hierarchical, or based on discrete components or “stages” in visual processing, but is also in line with a coarse-to-fine view of visual perception, in which the same neural populations require a longer accumulation of visual inputs for FI than (generic) face categorization (e.g., Sergent, 1986; Watt, 1987; Bachmann, 1991; Hegdé, 2008; Goffaux et al., 2011; Rossion, 2014).

It may be asked: to what extent are the measures of FI speed reported here *generalizable*? Here, a moderately sized sample of relatively homogenous, university students (age range: 19 to 28 years old) was used. This may have contributed to few inter-individual differences in behavior at lower stimulus viewing times (**Fig. S2 B&C**), and the relatively consistent stimulus viewing time limit for neural FI responses, this being 50 ms (20 Hz) for more than half the participants (**Fig. 6 A**). In future studies, processing time could also be compared within or across more diverse samples. For example, in cases of prosopagnosia, individuation of unfamiliar faces can often be performed well above chance level at long stimulus durations, when given enough time to apply feature-based strategies (Benton & Van

Allen, 1972; Davidoff & Landis, 1990), but not at short stimulus durations, when rapid, automatic FI is required (see Liu-Shuang, Torfs & Rossion, 2016). For another example, during development, children are able to individuate faces, but their performance is limited until adulthood (Carey, 1992). Better performance for upright than inverted faces emerges at a few years of age only and develops throughout childhood (Carey & Diamond, 1994; de Heering et al., 2012; Hills & Lewis, 2018). FI might thus require more time for children (e.g., six year-old children might not have a significant FI response at 50 ms), with more pronounced individual differences retained at relatively longer stimulus durations.

The present results are also likely influenced by stimulus factors. First, stimulus masking should be considered: the measures of FI speed taken here may be less conservative than if more optimized visual masks were used (Crouzet & Thorpe, 2011; Potter, 2012; Robinson, Grootswagers, & Carlson, 2019), although it may be argued that faces are ideal (high-level) masks for faces. Second, and most importantly, the facial identity stimuli used here certainly play a role: FI might occur at shorter stimulus durations if more variable facial identities were included, e.g., varying in gender, race, and age, yet might occur only with longer stimulus durations if images varied more extensively, e.g., were unsegmented, containing different viewpoints, lighting conditions, etc. Additionally, the variation in size and luminance at each stimulus presentation here, to reduce diagnostic low-level cues between base and oddball face stimuli, undoubtedly made FI more challenging.

In this context, an obviously important factor is the degree of familiarity of the participants with the morphological characteristics of the face set, and the specific face identities used in the experiment. Here, only Caucasian faces, i.e. of the same “race” as the participants, were used in the experiment, with participants most experienced with this race. Given the well-known other-race face effect observed across a variety of tasks and stimuli (Meissner & Brigham, 2001; Rossion & Michel, 2011 for review), using pictures of another, less familiar, race of faces (e.g., African or East Asian faces) in the same paradigm with the same participant race experience group may well lead to an increase in the minimal and optimal stimulus duration for FI. Conversely, participants with little everyday experience with Caucasian faces may show the opposite effect.

Moreover, while the present observations demonstrate once again that neurotypical human adults are very efficient at individuating pictures of unfamiliar faces, long-term familiarity with specific facial identities substantially increases generalization abilities of this individuation function (e.g., matching facial identities across views, Jenkins et al., 2011). Hence, even though the source of this increase in generalization remains unknown and may be

due to associations of variable views of faces with semantic, affective and verbal information (Dixon et al., 1998; Yovel & Schwartz, 2016; Rossion, 2018), the very same experiment performed with participants that are familiar with these faces (e.g., Verosky et al., 2020) may lead not only to a reduction of minimal and optimal stimulus duration values, but also to a reduction of inter-individual variability.

In the paradigm applied here, there are several factors that, while specific, should not limit the generalization of the findings. First, image predictability was not an important factor here, since an agreement between neural and behavioral measures in the minimal stimulus duration for FI was demonstrated, despite faces appearing periodically in the neural experiment and non-periodically in the behavioral experiment (see also Quek & Rossion, 2017). Additionally, such agreement across experiments also suggests that the task did not greatly affect these results, despite participants performing a fixation cross task in the first experiment, but explicitly individuating faces in the second experiment (see also Yan, Liu-Shuang & Rossion, 2019, for only late (post 180 ms) effects of task on the FI response). Third, while this paradigm relies on variable base to oddball stimulus presentation ratios, this is also not likely a limiting factor, since FI responses at 167 ms, 250 ms, and 250 ms (6 Hz, 5 Hz, and 4 Hz), showed stable amplitudes (**Fig. 4 A**; **Table 2 A**), despite varying ratios (see also Retter & Rossion, 2016).

### **The speed of face individuation**

In sum, we determined that within about 50 ms, the human visual system is able to accumulate the necessary visual inputs for FI, which may continue to be processed for additional time (differential, EEG FI responses lasted over 300 ms post oddball stimulus onset here). We also identified the optimal viewing time, i.e., the shortest duration providing the maximal FI response, at 167 ms at the neural level (as in Alonso-Prieto et al., 2013; Gentile & Rossion, 2014). While behavioral performance at 167 ms duration was at about 82% and continued to increase until the longest stimulus duration (333 ms), this improvement is likely to be due to general factors (e.g., increase in response confidence following change detection in the stimulation sequence). Indeed, contrary to mid-range frequencies, behavioral performance was unrelated to the magnitude of the neural FI response at the longest stimulus duration. Overall, these results support the view that a single glance suffices for neurotypical observers to fully capture the idiosyncratic features and configuration of an unfamiliar face. Such rapid and automatic high-level visual perception may be essential for our interaction with relevant stimuli in dynamic, visually complex environments.

Differences across individuals in their ability to individuate faces at middle stimulus durations negatively correlated with their negative, peak EEG latencies (occurring at ~ 185 ms on average), but not with an earlier, positive peak latency. This suggests that inter-individual differences in FI speed we measured behaviorally relate to high-level FI processes, rather than generic visual or anatomophysiological differences. Future studies could compare individuals' FI speed with other measures of FI ability, potentially also measuring the specificity of these FI effects against other neural and/or behavioral measures of face processing, e.g., generic face vs. object categorization or age or gender perception.

## **Acknowledgments**

We are grateful to Andrea Conte for programming the stimulation platform; to Renaud Laguesse, for creating the stimulus set; to Charles Or for generating the color key for plotting electrode locations on the two dimensional scalp map; to two anonymous reviewers; and to all members of the Face Categorization Lab who assisted in data collection. This research was supported by grants from the European Research Council [ERC: grant facessvep number 284025 to BR]; the Belgian National Foundation for Scientific Research [FNRS: grant numbers FC7159 to TR and PDR T.0207.16 FNRS to BR]; the National Eye Institute [<https://nei.nih.gov/>; grant number EY010834 to MW, and grant number EY023268 to FJ], and the Face perception INTER project (grant number INTER/FNRS/15/11015111 to CS) funded by the Luxembourgish Fund for Scientific Research (FNR, Luxembourg).

## References

- Alonso-Prieto, E.A, Van Belle, G., Liu-Shuang, J., Norcia, A.M., Rossion, B. (2013). The 6Hz fundamental frequency rate for individual face discrimination in the right occipito-temporal cortex. *Neuropsychologia*, 51, 2863-2975.
- Amihai, I., Deouell, L., & Bentin, S. (2011). Conscious awareness is necessary for processing race and gender information from faces. *Consciousness and cognition*, 20(2), 269–279.
- Bachmann, T. (1991). Identification of spatially quantised tachistoscopic images of faces: How many pixels does it take to carry identity? *European Journal of Cognitive Psychology*, 3(1), 87-103.
- Bacon-Macé, N., Macé, M. J., Fabre-Thorpe, M. & Thorpe, S. J. (2005). The time course of visual processing: backward masking and natural scene categorisation. *Vision Research*, 45(11), 1459-1469.
- Benton, A. L. & Van Allen, M. W. (1972). Prosopagnosia and facial discrimination. *Journal of the Neurological Sciences*, 15(2), 167-172.
- Besson, G., Barragan-Jason, G., Thorpe, S> J., Fabre-Thorpe, M., Puma, S., Ceccaldi, M. & Barbeau, E. J. (2017). From face processing to face recognition: Comparing three different processing levels. *Cognition*, 158, 33-43.
- Bindemann, M., Fysh, M., Cross, K., & Watts, R. (2016). Matching faces against the clock. *i-Perception*, 7(5), 2041669516672219, 1-18
- Bowles, D.C., McKone, E., Dawel, A., Duchaine, B., Palermo, R., Schmalzl, L., Rivolta, D., Wilson, E., & Yovel, G. (2009). Diagnosing prosopagnosia: effects of ageing, sex, and participant-stimulus ethnic match on the Cambridge Face Memory Test and Cambridge Face Perception Test. *Cognitive Neuropsychology*, 26, 423-455.
- Burton, A.M., White, D., & McNeill, A. (2010). The Glasgow Face Matching Test. *Behavior Research Methods*, 42, 286-291.
- Caharel, S., d'Arripe, O., Ramon, M., Jacques, C., Rossion, B. (2009). Early adaptation to unfamiliar faces across viewpoint changes in the right hemisphere: evidence from the N170 ERP component. *Neuropsychologia*, 47, 639-643.
- Carey, S. (1992). Becoming a face expert. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 335(1273), 95-103.
- Carey, S. & Diamond, R. (1994). Are faces perceived as configurations more by adults than by children? *Visual Cognition*, 1(2/3), 253-274.

- Crouzet, S. M., Kirchner, H. & Thorpe, S. J. (2010). Fast saccades toward faces: face detection in just 100 ms. *Journal of Vision*, 10:3.
- Crouzet, S. M. & Thorpe, S. J. (2011). Low-level cues and ultra-fast face detection. *Frontiers in Psychology*, 2:342.
- 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.
- Davidoff, J. & Landis, T. (1990). Recognition of unfamiliar faces in prosopagnosia. *Neuropsychologia*, 28(11), 1143-1161.
- de Heering, A., Rossion, B., Maurer, D. (2012). Developmental changes in face recognition during childhood: evidence from upright and inverted faces. *Cognitive Development*, 27, 17-27.
- Dobs, K., Isik, L., Pantazis, D., & Kanwisher, N. (2019). How face perception unfolds over time. *Nature Communications*, 10(1), 1258.
- Dixon, M. J., Bub, D. N. & Arguin, M. (1998). Semantic and visual determinants of face recognition in a prosopagnosic patient. *Journal of Cognitive Neuroscience*, 10(3), 362-376.
- Duchaine B. & Nakayama, K. (2006). The Cambridge Face Memory Test: Results for neurologically intact individuals and an investigation of its validity using inverted face stimuli and prosopagnosic participants. *Neuropsychologia*, 44(4), 576-585.
- Dzhelyova, M. & Rossion, B. (2014). The effect of parametric stimulus size variation on individual face discrimination indexed by fast periodic visual stimulation. *BMC Neuroscience*, 15:87, 1-12.
- Dzhelyova, M. Schiltz, C., & Rossion, B. (2020). The relationship between the Benton Face Recognition Test and electrophysiological unfamiliar face individuation response as revealed by fast periodic visual stimulation. *Perception*, 49(2), 210–221.
- Foldiak, P., Xiao, D., Keysers, C., Edwards, R. & Perrett, D. I. (2004). Rapid serial visual presentation for the determination of neural selectivity in area STSa. *Progress in Brain Research*, 144, 107-116.
- Fysh, M. C., & Bindemann, M. (2017). Effects of time pressure and time passage on face-matching accuracy. *Royal Society Open Science*, 4(6):170249.
- Gabbiani, F. & Cox, S. (2017). *Mathematics for Neuroscientists*. Academic Press.

- Ganis, G., Smith, D. & Schendan, H. E. (2012). The N170, not the P1, indexes the earliest time for categorical perception of faces, regardless of interstimulus variance. *NeuroImage*, 62(3), 1563-1574.
- Gao, X., Gentile, F. & Rossion, B. (2018). Fast periodic stimulation: a highly effective approach in fMRI brain mapping. *Brain Structure and Function*, 223, 2433-2454.
- Gentile, F., & Rossion, B. (2014). Temporal frequency tuning of cortical face-sensitive areas for individual face perception. *NeuroImage*, 90, 256-265.
- Goffaux, V., Peters, J., Haubrechts, J., Schiltz, C. Jansma, B., Goebel, R. (2011). From coarse to fine? Spatial and temporal dynamics of cortical face processing. *Cerebral Cortex*, 21(2), 467-476.
- Grill-Spector, K., Kushnir, T., Hendler, T. & Malach, R. (2000). The dynamics of object-selective activation correlate with recognition performance in humans. *Nature Neuroscience*, 3(8), 837-843.
- Grill-Spector, K. & Malach, R. (2001). fMR-adaptation: a tool for studying the functional properties of human cortical neurons. *Acta Psychologica*, 107(1-3), 293-321.
- Grill-Spector, K., Henson, R., & Martin, A. (2006). Repetition and the brain: neural models of stimulus-specific effects. *Trends in Cognitive Sciences*, 10, 14-23.
- Grill-Spector, K., Weiner, K. S. (2014). Functional architecture of human ventral temporal cortex and its role in categorization. *Nature Reviews Neuroscience*, 15(8), 536-548.
- Gur, M. (2018). Very small faces are easily discriminated under long and short exposure times. *Journal of Neurophysiology*, 119(5), 1559-1607.
- Hegd , J., Fang, F., Murray, S. O. & Kersten, D. (2008). Preferential responses to occluded objects in the human visual cortex. *Journal of Vision*, 8(4):16.
- Heisz, J. J., Watter, S. & Shedden, J. M. (2006). Progressive N170 habituation to unattended repeated faces. *Vision Research*, 46(1-2), 47-56.
- Henson, R. N. (2016). Repetition suppression to faces in the fusiform face area: A personal and dynamic journey. *Cortex*, 80, 174-184.
- Hermann, P., Grotheer, M., Kovacs, G. & Vidnyanszky, Z. (2017). The relationship between repetition suppression and face perception. *Brain, Imaging and Behavior*, 11(4), 1018-1028.
- Herzmann, G., Kunina, O., Sommer, W. & Wilhelm, O. (2009). Individual differences in face cognition: brain-behavior relationships. *Journal of Cognitive Neuroscience*, 22:3, 571-589.
- Herzmann, G, Kunina, O., Sommer, W., & Wilhelm, O. (2010). Individual differences in face

- cognition: brain-behavior relationships. *Journal of Cognitive Neuroscience*, 22, 571-589.
- Hildebrandt, A., Wilhelm, O., Schmiedek, F., Herzmann, G. & Sommer, W. (2011). On the specificity of face cognition compared with general cognitive functioning across adult age. *Psychology and Aging*, 26, 701–715.
- Hills, P. J., & Lewis, M. B. (2018). The development of face expertise: Evidence for a qualitative change in processing. *Cognitive Development*, 48, 1–18.
- Hsiao, J. H., & Cottrell, G. (2008). Two fixations suffice in face recognition. *Psychological Science*, 19, 998-1006.
- Hubel, D. H. & Wiesel, T. N. (2005). *Brain and visual perception: The story of a 25-year collaboration*. Oxford University Press.
- Jacques, C. & Rossion, B. (2006). The speed of individual face categorization. *Psychological Science*, 17, 485-492.
- Jacques, C. & Rossion, B. (2007). Early electrophysiological responses to multiple face orientations correlate with individual discrimination performance in humans. *NeuroImage*, 36, 863-876.
- Jenkins, R., White, D., Van Montfort, X. & Burton, A. M. (2011). Variability in photos of the same face. *Cognition*, 121(3), 313-323.
- Kaltwasser, L., Hildebrandt, A., Recio, G., Wilhelm, O. & Sommer, W. (2014). Neurocognitive mechanisms of individual differences in face cognition: a replication and extension. *Cognitive, Affective Behavioral Neuroscience*, 14, 61-78.
- Keysers, C., Xiao, D.- K., Földiák, P., & Perrett, D. I. (2001). The speed of sight. *Journal of Cognitive Neuroscience*, 13, 90–101.
- Keysers, C. & Perrett, D. I. (2002). Visual masking and RSVP reveal neural competition. *Trends in Cognitive Science*, 6, 120–125.
- Keysers, C., Xiao, D.-K., Foldiak, P. & Perrett, D. I. (2005). Out of sight but not out of mind: the neurophysiology of iconic memory in the superior temporal sulcus. *Cognitive Neuropsychology*, 22, 3-4, 316-332.
- Kovács, G. Vogels, R. & Orban, G. A. (1995). Cortical correlate of pattern backward masking. *PNAS USA*, 92(12), 5587-5591.
- Kravitz, D. J., Saleem, K. S., Baker, C. I., Ungerleider, L. G. & Mishkin, M. (2013). The ventral visual pathway: an expanded neural framework for the processing of object quality. *Trends in Cognitive Sciences*, 17(1), 26–49.

- Laguesse, R., Dormal, G., Biervoye, A., Kuefner, D. & Rossion, B. (2012). Extensive visual training in adulthood significantly reduces the face inversion effect. *Journal of Vision*, 12(14), 1-13.
- Lehky, S. R. (2000). Fine discrimination of faces can be performed rapidly. *Journal of Cognitive Neuroscience*, 12:5, 848-855.
- 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 visual 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.
- Lueschow, A., Weber, J. E., Carbon, C.-C., Deffke, I., Sander, T., Gruter, T., Gruter, M., Trahms, L. & Curio, G. (2015). The 170ms response to faces as measured by MEG (M170) is consistently altered in congenital prosopagnosia. *PLOS One*, 0137624.
- McCaffery, J. M., Robertson, D. J., Young, A. W., & Burton & A. W. (2018). Individual differences in face identity processing. *Cognitive Research: Principles and Implications*, 3:21.
- McCarthy, G., & Wood, C. C. (1985). Scalp distributions of event-related potentials: An ambiguity associated with analysis of variance models. *Electroencephalography and Clinical Neurophysiology, Evoked Potentials Section*, 62, 203–208.
- McGugin, R. W., Ryan, K. F., Tamber-Rosenau, B. J. & Gauthier, I. (2018). The role of experience in the face-selective response in right FFA. *Cerebral Cortex*, 28(6), 2071-2084.
- Meissner, C. A., & Brigham, J. C. (2001). Thirty years of investigating the own-race bias in memory for faces: A meta-analytic review. *Psychology, Public Policy, and Law*, 7(1), 3–35.
- Mohsenzadeh, Y., Qin, S., Cichy, R. M. & Pantazis, D. (2018). Ultra-Rapid serial visual presentation reveals dynamics of feedforward and feedback processes in the ventral visual pathway. *eLife*, 7:e36329.
- Nasanen, R., Ojanpaa, H., Tanskanen, T. & Paallysaho, J. (2006). Estimation of temporal resolution of object identification in human vision. *Experimental Brain Research*, 172, 464-471.

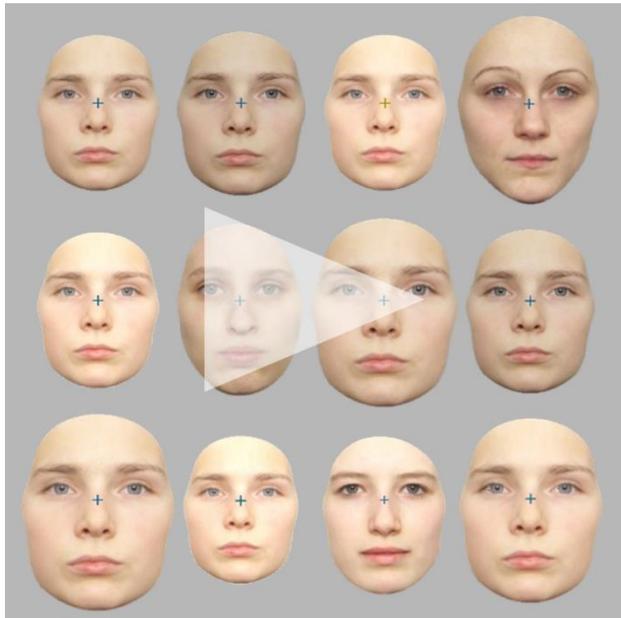
- Nemrodov, D., Jacques, C., Rossion, B. (2015). Temporal dynamics of repetition suppression to individual faces presented at a fast periodic rate. *International Journal of Psychophysiology*, 98, 35-43.
- Nemrodov, D., Niemeier, M., Patel, A. & Nestor, A. (2016). The neural dynamics of facial identity processing: insights from EEG-based pattern analysis and image reconstruction. *eNeuro*, 0358-17.2018.
- Nemrodov, D., Ling, S., Nudnou, I., Roberts, T., Cant, J. S., Lee, A. C. H. & Nestor, A. (2019). A multivariate investigation of visual word, face, and ensemble processing: Perspectives from EEG - based decoding and feature selection. *Psychophysiology*, 57(3), e13511.
- Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia*. 9(1), 97–113.
- Or, C. C.-F. & Wilson, H. R. (2010). Face recognition: Are viewpoint and identity processed after face detection? *Vision Research*, 50(16), 1581-1589.
- Orban de Xivry, J.-J., Ramon, M., Lefevre, P. & Rossion, B. (2008). Reduced fixation on the upper area of personally familiar faces following acquired prosopagnosia. *Journal of Neuropsychology*, 2, 245-268.
- Parketny, J., Towler, J. & Eimer, M. (2015). The activation of visual face memory and explicit face recognition are delayed in developmental prosopagnosia. *Neuropsychologia*, 75, 538-547.
- Peterson, M. F. & Eckstein, M. P. (2013). Individual differences in eye movements during face identification reflect observer-specific optimal points of fixation. *Psychological Science*, 24(7), 1216-1225.
- Potter, M. C. (2012). Recognition and memory for briefly presented scenes. *Frontiers in Psychology*, 3:32.
- Quek, G. & Rossion, B. (2017). Category-selective human brain processes elicited in fast periodic visual stimulation streams are immune to temporal predictability. *Neuropsychologia*, 104, 182-200.
- Ramon, M., Caharel, S., & Rossion, B. (2011). The speed of personally familiar face recognition. *Perception*, 40, 437-449.
- Retter, T. L. & Rossion, B. (2016). Uncovering the neural magnitude and spatio-temporal dynamics of natural image categorization in a fast visual stream. *Neuropsychologia*, 91, 9-28.

- Retter, T. L., Jiang, F., Webster, M. A. & Rossion, B. (2018). Dissociable effects of inter-stimulus interval and presentation duration on rapid face categorization. *Vision Research*, 145, 11-20.
- Retter, T. L., Jiang, F., Webster, M.A., Rossion, B. (2020). All-or-none visual categorization in the human brain. *NeuroImage*, 213:116685, 1-16.
- Retter, T. L., Rossion, B. & Schiltz, C. (2021). Harmonic amplitude summation for frequency-tagging analysis. *Zenodo*. <https://doi.org/10.5281/zenodo.4600949>
- Robinson, A. K., Grootswagers, T. & Carlson, T. A. (2019). The influence of image masking on object representations during rapid serial visual presentation. *NeuroImage*, 197, 224-231.
- Rolls, E. T., Tovee, M. J., Purcell, D. G., Stewart, A. L. & Azzopardi, P. (1994). The responses of neurons in the temporal cortex of primates, and face identification and detection. *Experimental Brain Research*, 101, 473-484.
- Rossion, B. (2014). Understanding individual face discrimination by means of fast periodic visual stimulation. *Experimental Brain Research*, 232, 1599-1621.
- Rossion, B. (2018). Humans are visual experts at unfamiliar face recognition. *Trends in Cognitive Sciences*, 22, 471-472.
- Rossion, B. & Boremanse, A. (2011). Robust sensitivity to facial identity in the right human occipito-temporal cortex as revealed by steady-state visual-evoked potentials. *Journal of Vision*. 11(2):16, 1–21.
- Rossion, B. & Caharel, S. (2011). ERP evidence for the speed of face categorization in the human brain: disentangling the contribution of low-level visual cues from face perception. *Vision Research*, 51, 1297-1311.
- Rossion, B. & Jacques, C. (2011). The N170 : understanding the time-course of face perception in the human brain. *The Oxford Handbook of ERP Components*, Edited by S. Luck and E. Kappenman. Oxford University Press.
- Rossion, B., Torfs, K., Jacques, C., Liu-Shuang, J. (2015). Fast periodic presentation of natural face images reveals a robust face-selective electrophysiological response in the human brain. *Journal of Vision*, 15(1), 1-18.
- Rossion, B. & Michel, C. (2018). Normative accuracy and response time data for the computerized Benton Facial Recognition Test (BFRT-c). *Behavior Research Methods*, 50, 2442-2460.

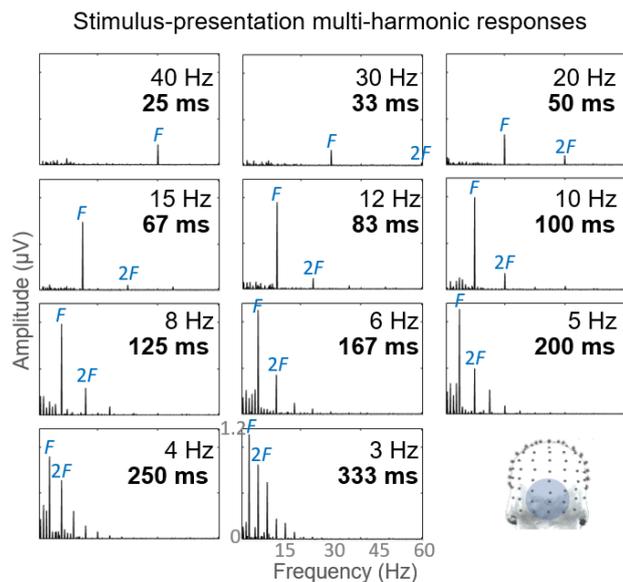
- Rossion, B., Retter, T. L. & Liu-Shuang, J. (2020). Understanding human individuation of unfamiliar faces with oddball fast periodic visual stimulation and electroencephalography. *European Journal of Neuroscience*, 52, 4283-4334.
- Schwartz, L. & Yovel, G. (2016). The roles of perceptual and conceptual information in face recognition. *Journal of Experimental Psychology. General*, 145(11), 1493–1511.
- Schweinberger, S. R. & Neumann, M. F. (2016). Repetition effects in human ERPs to faces. *Cortex*, 80, 141-153.
- Seeck, M., Michel, C. M., Mainwaring, N., Cosgrove, R., Blume, H., Ives, J., Landis, T. & Schomer, D. L. (1997). Evidence for rapid face recognition from human scalp and intracranial electrodes. *Cognitive Neuroscience and Neuropsychology*, 8(12), 2749-2754.
- Sergent, J. (1986). Microgenesis of Face Perception. In: H.D. Ellis, M. A. Jeeves, F. Newcombe & A. Young (Eds), *Aspects of Face Processing*. NATO ASI Series (Series D: Behavioural and Social Sciences), vol 28. Springer, Dordrecht.
- Stacchi, L., Liu-Shuang, J., Ramon, M. & Caldara, R. (2019). Reliability of individual differences in neural face identity discrimination. *Neuroimage*, 189, 468-475.
- Strasburger, H., Scheidler, W. & Rentschler, I. (1988). Amplitude and phase characteristics of the steady-state visual evoked potential. *Applied Optics*, 27(6), 1069-1088.
- Su, J., Chen, C., He, D. & Fang, F. (2012). Effects of face view discrimination learning on N170 latency and amplitude. *Vision Research*, 61, 125-131.
- Tanskanen, T., Näsänen, R., Ojanpää, H. & Hari, R. (2007). Face recognition and cortical responses: Effect of stimulus duration. *NeuroImage*, 35, 1636-1644.
- Thorpe, S., Fize, D. & Marlot, C. (1996). Speed of processing in the human visual system. *Nature*, 381(6582), 520-522.
- Towler, J., Gosling, A., Duchaine, B. & Eimer, M. (2012). The face-sensitive N170 component in developmental prosopagnosia. *Neuropsychologia*, 50(14), 3588-3599.
- Towler, J., Fisher, K. & Eimer, M. (2017). The cognitive and neural basis of developmental prosopagnosia. *The Quarterly Journal of Experimental Psychology*, 70(2), 316-344.
- Townsend, J. T. & Ashby, F. G. (1983). *The Stochastic Modeling of Elementary Psychological Processes*. Cambridge: Cambridge University Press.
- Van den Berg, B., Appelbaum, L. W., Clark, K., Lorist, M. M. & Woldorff, M. G. (2016). Visual search performance is predicted by both prestimulus and poststimulus electrical brain activity. *Scientific Reports*, 6:37718, 1-13.

- Vanni, S., Revonsuo, A., Saarinen, J. & Hari, R. (1996). Visual awareness of objects correlates with activity of right occipital cortex. *Neuroreport*, 8(1), 183-186.
- Verosky, S. C., Zoner, K. A., Marble, C. W., Sammon, M. M., & Babarinsa, C. O. (2020). Familiarization increases face individuation measured with fast periodic visual stimulation. *Biological psychology*, 153, 107883.
- Watt, R. J. (1987). Scanning from coarse to fine spatial scales in the human visual system after the onset of a stimulus. *Journal of the Optical Society of America A*, 4, 2006-2021.
- Wu, Q., Yan, B., Zeng, Y., Zhang, C. & Tong, L. (2018). Anti-deception: reliable EEG-based biometrics with real-time capability from the neural response of face rapid serial visual presentation. *BioMedical Engineering OnLine*, 17:55, 1-16.
- Xu, B., Liu-Shuang, J., Rossion, B. & Tanaka, J. (2017). Individual differences in face identity processing with fast periodic visual stimulation. *Journal of Cognitive Neuroscience*, 29(8), 1368-1377.
- Yan, X., Liu-Shuang, J., & Rossion B. (2019). Effect of face-related task on rapid individual face discrimination. *Neuropsychologia*, 129, 236-245.
- Zheng, X., Mondloch, C. J. & Segalowitz, S. J. (2011). The timing of individual face recognition in the brain. *Neuropsychologia*, 50(7), 1451-1461.

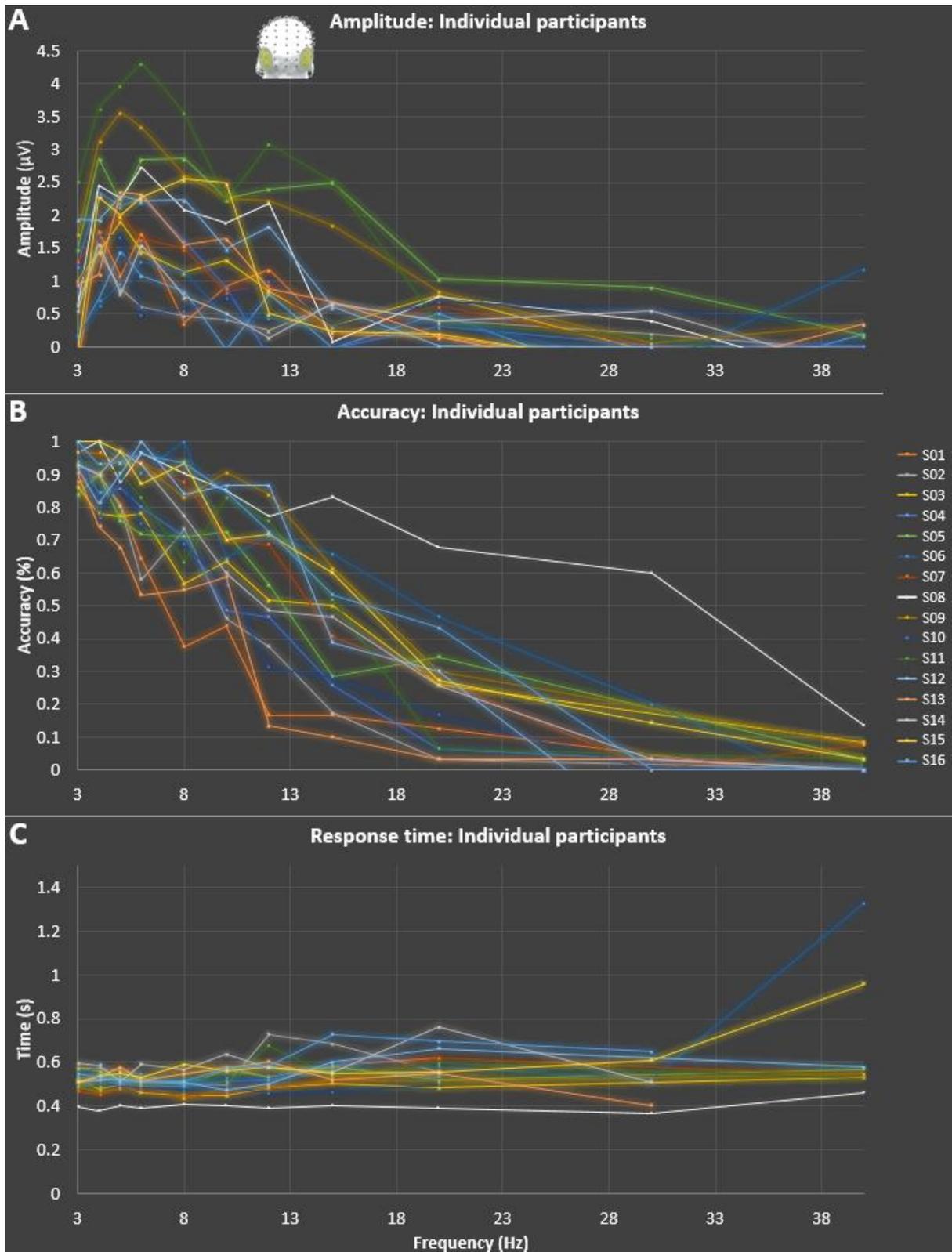
## Supplemental Material



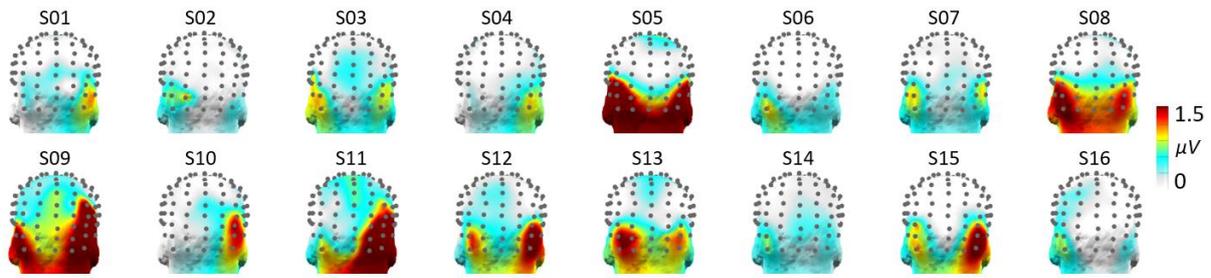
**Movie S1.** A shortened example experimental trial, consisting of the nine 7-s stimulus duration steps of  $F$  available for display on a common 60 Hz monitor (33, 50, 67, 83, 100, 167, 200, 250, and 333 ms; i.e., 30, 20, 15, 12, 10, 6, 5, 4, and 3 Hz); excluding 25 and 125 ms (40 and 8 Hz)). Throughout, a facial identity changes occurs at  $F/n = 1$  Hz (1 s).



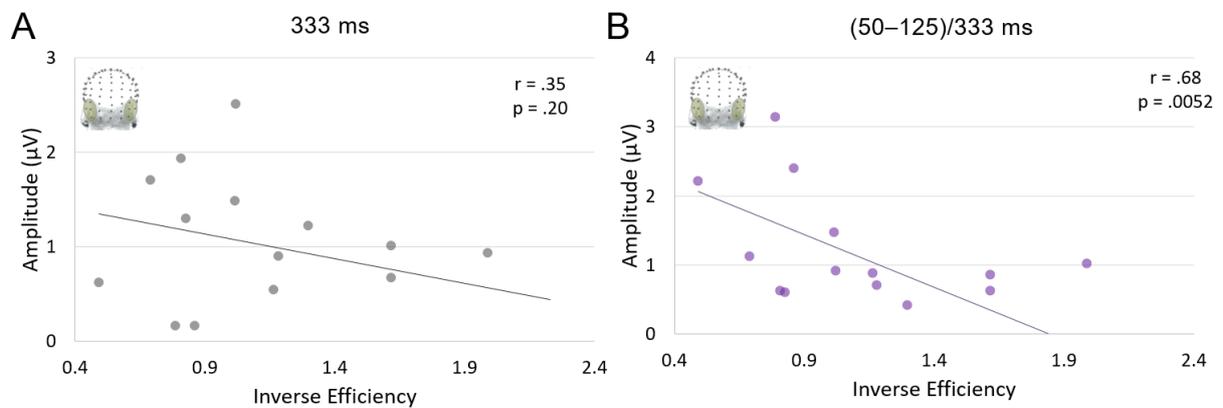
**Figure S1.** Frequency-domain, baseline-subtracted amplitude spectra for harmonic frequency responses to stimulus-presentation at  $F$ , defined by the stimulus presentation rate, and its harmonics up to 60 Hz, are viewed over the medial-occipital ROI.



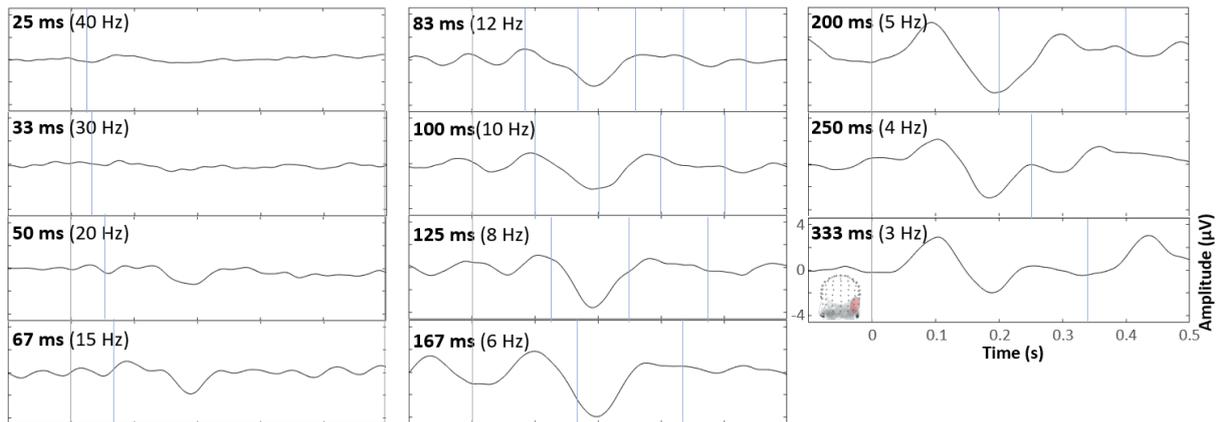
**Figure S2.** Individual participant face individuation behavior across presentation rates: each participant is drawn in a specific color as indicated, the same in each panel. **A)** Face individuation EEG amplitude over the bilateral occipito-temporal ROI. **B)** Accuracy (corrected for false positives). **C)** Correct response time.



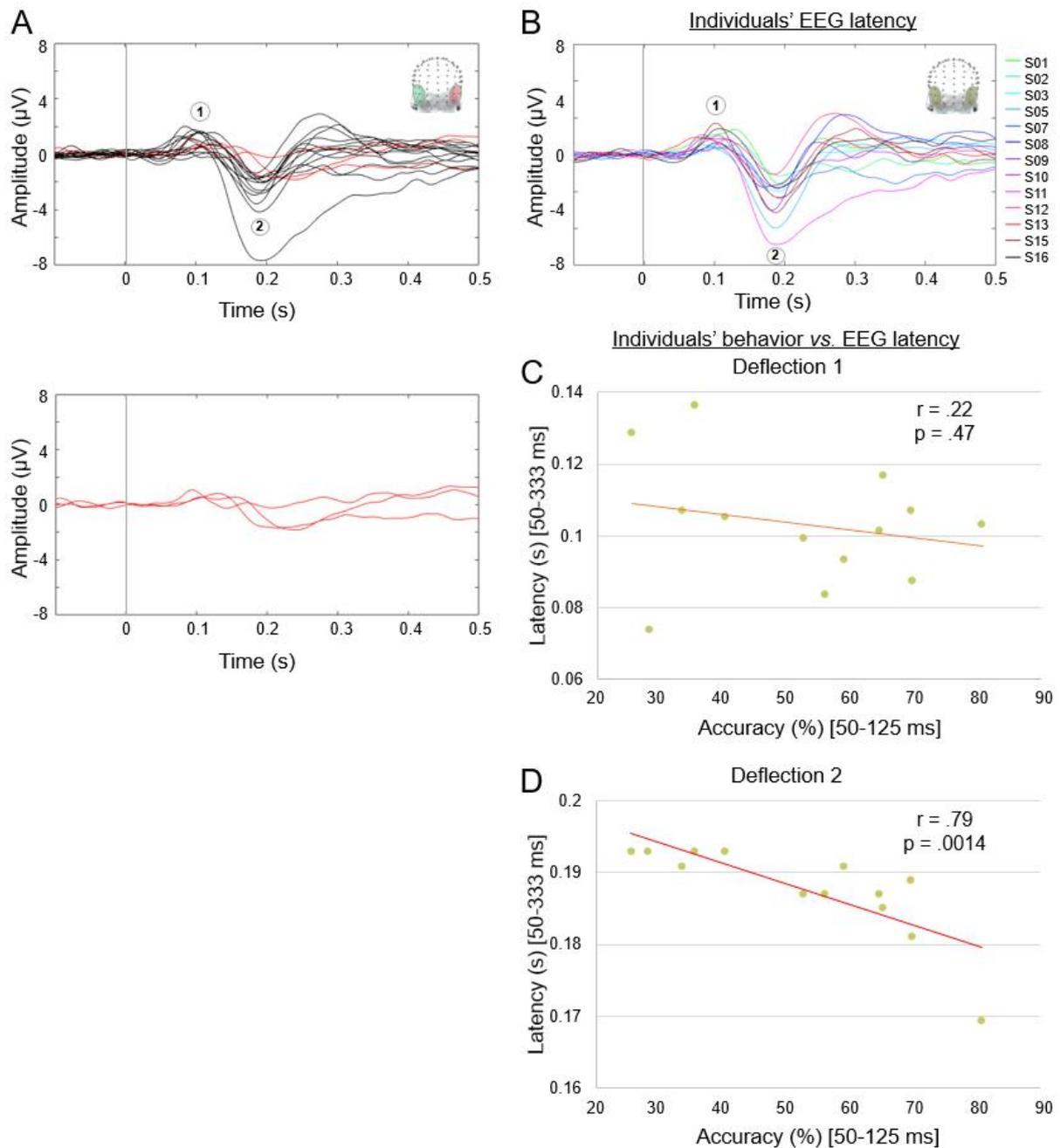
**Figure S3.** Individual participant, harmonic-summed, baseline-subtracted face individuation scalp topographies, averaged across a mid-range of stimulus-presentation rates, from 50 to 125 ms (20 to 8 Hz), plotted to a common amplitude scale.



**Figure S4.** Correlations between individual participant face individuation inverse efficiency, i.e., correct response time (s) / accuracy (proportion correct), measured in the behavioral experiment from 50 to 125 ms (20 to 8 Hz), and frequency-domain amplitude from the EEG frequency-tagging experiment (stimulus duration as indicated; base-subtracted, summed harmonics of 1 Hz, over the OT ROI), corresponding to **Fig. 5** for accuracy. **A**) No significant correlation of individuals' inverse efficiency with amplitude at the longest stimulus duration condition, 333 ms (3 Hz). **B**) A significant correlation of individuals' inverse efficiency with amplitude in the middle stimulus duration range, 50 to 125 ms (20 to 8 Hz), weighted by the amplitude at 333 ms (3 Hz).



**Figure S5.** Time-domain face individuation EEG responses shown as in **Fig. 7**, except that the responses general to visual stimulation have not been removed by a frequency-domain notch filter. A blue bar indicates the time that the next “base” facial identity was presented following the “oddball” identity presentation, which occurred continuously throughout the stimulation (e.g., there was one additional base face presented in this time window in the 167 ms duration (6 Hz) condition, two additional base faces in the 125 ms duration (8 Hz) condition, etc.). Note that in pre-processing for the time domain, the data were low-pass filtered at 30 Hz, reducing stimulus-presentation rate responses at the 33 and 25 ms (30 and 40 Hz) conditions in this display (see **Fig. 4 D** for these response amplitudes in the frequency domain).



**Figure S6.** *A) Removal of participants with poor time domain responses in the EEG frequency-tagging experiment (50 to 333 ms; 20 to 3 Hz). Upper panel: face individuation responses from three participants not included in the main analyses are drawn in red, together with responses from all included participants drawn in black. Lower panel: These three removed participants' responses are shown in isolation. These participants' (S04, S06, and S14) low-amplitude frequency-domain face individuation topographies can be found in Fig. S3. Data were plotted at either the right or left occipito-temporal sub-region, corresponding to the lateralization index of each participant from the frequency-domain analyses. B) Similar responses were produced in the time domain when investigating individual participants'*

*bilateral occipito-temporal responses (as in Fig. 8, showing right or left sub-region responses). C) For the first deflection, there was an insignificant, weak, negative correlation between participants' behavioral accuracy and the peak latency, slope = -0.21 ms/%,  $r = .22$ ,  $p = .47$ . D) For the second deflection, there was a significant, negative correlation, slope = -0.29 ms/%,  $r = .79$ ,  $p = .0014$ .*