The 6 Hz fundamental stimulation frequency rate for individual face discrimination in the right occipito-temporal cortex

Esther Alonso-Prieto¹,a, Goedele Van Bellea, Joan Liu-Shuanga, Anthony M. Norcia¹,b, Bruno Rossiona,*

¹Institute of Research in Psychology and Institute of Neuroscience, University of Louvain, 10 Place Cardinal Mercier, Louvain-la-Neuve 1348, Belgium
²Department of Psychology, Stanford University, 450 Serra Mall, Stanford, CA 94305, USA

A R T I C L E   I N F O

Article history:
Received 2 June 2013
Received in revised form
29 July 2013
Accepted 26 August 2013
Available online 2 September 2013

Keywords:
Face perception
Frequency tuning
SSVEP
6 Hz
Theta range N170
Identity adaptation

A B S T R A C T

What is the stimulus presentation rate at which the human brain can discriminate each exemplar of a familiar visual category? We presented faces at 14 frequency rates (1.0–16.66 Hz) to human observers while recording high-density electroencephalogram (EEG). Different face exemplars elicited a larger steady-state visual evoked (ssVEP) response than when the same face was repeated, but only for stimulation frequencies between 4 and 8.33 Hz, with a maximal difference at 5.88 Hz (170 ms cycle). The effect was confined to the exact stimulation frequency and localized over the right occipito-temporal cortex. At high frequency rates (> 10 Hz), the response to different and identical exemplars did not differ, suggesting that the fine-grained analysis needed for individual face discrimination cannot be completed before the next face interrupts, or competes, with the processed face. At low rates (< 3 Hz), repetition suppression could not be identified at the stimulation frequency, suggesting that the neural response to an individual face is temporally dispersed and distributed over different processes. These observations indicate that at a temporal rate of 170 ms (6 faces/s) the face perception network is able to fully discriminate between each individual face presented, providing information about the temporal bottleneck of individual face discrimination in humans. These results also have important practical implications for optimizing paradigms that rely on repetition suppression, and open an avenue for investigating complex visual processes at an optimal range of stimulation frequency rates.

© 2013 Elsevier Ltd. All rights reserved.

1. Introduction

The visual system is able to extract diagnostic information not only at a coarse level to categorize a visual stimulus as a face ("face detection", e.g., (Crouzet, Kirchner, & Thorpe, 2010)) but also at a finer-grained level in order to discriminate it from other individual faces ("individual face discrimination", or more simply "face discrimination"). Behaviorally, human observers are able to discriminate different individual faces accurately in a few hundreds of milliseconds (e.g., (Jacques, d’Arripe, & Rossion, 2007)). Face-selective cells of the same neuronal population in the monkey infero-temporal (IT) cortex (Gross, Rocha-Miranda, & Bender, 1972) discharge at different rates to different individual faces, suggesting a mechanism to discriminate faces based on sparse population coding (Abbott, Rolls, & Tovee, 1996; Leopold, Bondar, & Giese, 2006; Young & Yamane, 1992). Yet, in humans, neuroimaging studies have identified several areas of the ventral occipito-temporal cortex that are sensitive to differences between individual faces, showing a larger neural response when different faces are presented successively compared to the repetition of the exact same face ("repetition suppression", or "neural adaptation", e.g., (Davies-Thompson, Gouws, & Andrews, 2009; Gilae-Dotan, Gelbard-Sagiv, & Malach, 2010; Grill-Spector & Malach, 2001)).

An unexplored issue concerns the sensitivity of the human visual system to stimulus presentation rate. That is, how many faces can be discriminated in 1 s of time? This issue is different than defining the speed of conduction of information about faces, information that is usually inferred from either the latency of face-related event-related potentials (ERPs, e.g., (Bentin, Allison, Puce, Perez, & McCarthy, 1996; Jeffreys, 1996b; Rossion & Jacques, 2011; VanRullen, 2011)) or the latency of discharges of face-selective neurons in the monkey infero-temporal cortex (e.g. (Kiani, Esteky, & Tanaka, 2005; Sugase, Yamane, Ueno, & Kawano, 1999)). It is also a different issue than the presentation duration that is needed to recognize a single individual face and activate its neural representation (Tanskanen, Nasanen, Ojanpaa, & Hari, 2007). Rather, it concerns the time that is necessary to process a face at the individual level before the next one can be handled. This is an
important issue to resolve for at least four reasons. First, generally speaking, the temporal interval over which the system blurs information together, known as its ‘temporal resolution’, is important for understanding human visual perception and may be a fundamental attribute of brain function (Brown, 1965; Hawken, Shapley, & Grosof, 1996; Holcombe, 2009; Keysers, Xiao, Foldiak, & Perrett, 2001; Krukowski & Miller, 2001; Tovee, Rolls, & Azzopardi, 1994). Second, more specifically, humans live in a highly dynamic visual world in which they can be exposed to many different faces simultaneously, or within a short timeframe. Being able to discriminate these faces rapidly may be critical for adequate social interactions. Third, clarifying the maximal rate at which individual faces can be discriminated would have important implications for understanding the neuro-functional basis of face perception, putting constraints on the nature of the information that can be extracted during a certain amount of time and potentially transferred to higher levels of processing. Fourth, determining the frequency (rate) tuning function of the discrimination of individual faces would have practical implications for optimizing studies that rely on neural repetition suppression in high-level vision (Grill-Spector, Henson, & Martin, 2006).

The steady state visual evoked potential (ssVEP) technique recorded by means of electroencephalogram (EEG) or magnetoencephalogram (MEG) is well suited to address the above mentioned issue because it benefits from the observation that in humans (and monkeys, see e.g., (Nakayama & Mackeben, 1982)), a periodic stimulation elicits a periodic EEG/MEG response at the exact frequency of stimulation and its harmonics, the ssVEPs (Regan, 1966, 1989). Thus, it can be used to characterize the visual system’s sensitivity to stimulus presentation rate noninvasively, at a macroscopic brain level, with high signal-to-noise ratio and high frequency resolution (Regan, 1966, 1989). Recently, the ssVEP approach was extended to study face detection (Ales, Farzin, Rosson, & Norcia, 2012) and individual face discrimination (Rossion & Boremanse, 2011; Rossion, Prieto, Boremanse, Kuefner, & Van Belle, 2012), using single stimulation frequencies. In line with these latter studies, the temporal frequency tuning function for individual face discrimination was assessed here by comparing the electrophysiological response when individual faces changed at every cycle to the response when the same face was presented at every cycle.

Specifically, we measured neural temporal frequency tuning for face discrimination by showing human observers pictures of faces presented at different frequency rates (1 to 16.66 Hz, i.e., a face every 1000 to every ~60 ms; Fig. 1), while recording high-density scalp EEG. This range of frequencies was selected by considering several factors. First, a minimal stimulation frequency of 1 Hz is sufficiently low to clearly observe the transient ERP responses following face stimulation, given that the return to baseline in typical ERP studies usually takes about half a second (Nunez & Srivivasan, 2006). Second, the maximal responses to low-level visual stimulation in EEG/MEG studies are observed either at about 15 Hz (Hermann, 2001; Pastor, Artieda, Arbizu, Valencia, & Masdeu, 2003) or below (e.g., 8–10 Hz in (Fawcett, Barnes, Hillebrand, & Singh, 2004; Regan, 1966, 1989; Singh, Kim, & Kim, 2003; Srivivasan, Bibi, & Nunez, 2006; Van Der Tweel & Lunel, 1965)). Given that the rate of discrimination should be lower for high-level complex visual stimuli such as faces than for low-level visual stimuli, a range between 1 and 16.66 Hz should cover most if not all of the responses of interest. Finally, intermediate frequency rates were selected in order to be able to define a meaningful frequency tuning function for individual face discrimination. For this reason, we used a large number of frequency rates (12) which were relatively equally spaced, considering the technical constraints (refresh rate of the monitor). We also included one condition encompassing an exact cycle of 170 ms (i.e., 5.88 Hz) because a cycle of 170 ms corresponds to the peak of the face-sensitive N170 component (Bentin et al., 1996; Rossion & Jacques, 2008). Most importantly, the earliest repetition effects for individual faces are observed at the peak of the N170 (e.g., Itier & Taylor, 2002; Jacobs et al., 2007; Rossion & Jacques, 2011) and these effects are often prolonged until 250–300 ms (e.g., Caharel, d’Arripe, Ramon, Jacques, & Rossion, 2009; Schweinberger, Pickering, Jentsch, Burton, & Kaufmann, 2002). Even though the latency of ERPs does not reflect the temporal rate of processing, these observations suggest that a range of stimulation rates between 3.5/4 and 6 Hz might be associated with the largest difference between repeated and different individual faces.

2 Peaks of EEG responses at the stimulation frequency have been observed at higher frequency bands for sinusoidally modulated light stimulation (i.e., the medium, about 20 Hz, and high, about 40 Hz, frequency regions described by (Regan, 1966, 1989)). However, the medium and high frequency responses have not been described for patterned (structured) visual stimulation (Regan, 1989; Silberstein, Ciocca, & Phippas, 1995).

2. Materials and methods

2.1. Participants

Four healthy adult participants (ages 28, 30, 35, 36 years), all right-handed males with normal or corrected-to-normal vision, took part in the study for payment. They were tested four times on different days, over a period of four weeks (16 EEG recordings in total). Eight new right-handed participants (two males) were tested in a single EEG recording session for a complementary experiment. Written informed consent was obtained from all participants prior to the experiments, which were approved by the Biomedical Ethical Committee of the University of Louvain.
2.2. Procedure

The main aspects of the procedure were described in two recent EEG studies that compared different face trials to identical face trials when stimulated at a fixed frequency rate (3.5 Hz with laser-scanned color faces in (Rossion & Boremanse, 2011), and 4.0 Hz with grayscale faces in (Rossion et al., 2012)) and will be summarized here.

2.2.1. Stimuli

Eighteen full-front color pictures of faces from the Tubingen Max Planck Institute (MPI) database of laser-scanned (Cyberware TM) human heads were used. They were cropped to remove external features (hair and ears) but their overall shape was preserved. All face stimuli were unfamiliar to the participants and had a face area (100% size) subtended 3.72 ± 2.86 degrees of visual angle at a distance of 100 cm from the stimulation monitor. The stimuli were equalized for mean luminance online by the stimulation program.

2.2.2. Stimulation

After electrode-cap placement, participants were seated in a light- and sound-attenuated room. For each participant the viewing distance was measured and set to 100 cm from a computer monitor before the experiment started. There was no chin-rest, but instructions to maintain constant distance were provided to the participants and controlled regularly by the experimenter. Stimuli were displayed on a grey background (40 cd/m²) via a custom application (SinStim) running in Matlab (The Mathworks Inc., Natick, Mass.). In the ‘identical face’ condition, the same color face, chosen randomly from 18 faces, was presented repeatedly for 84 s (one trial; Fig. 1a). In the ‘different faces’ condition, the same face picture was presented for the first 15 s, then a different face than the previous one was presented at every cycle, until the end of the sequence (see (Rossion et al., 2012)). Since there were only 18 faces, the different face identities were also repeated in the ‘different faces’ condition, but with the important constraint that the same face identity could not be presented in consecutive cycles.

In each trial regardless of condition, a face stimulus appeared on the screen at a stimulation rate of F faces/s (one face every 1000/F ms, with 2F changes/s when considering both face onset and face disappearance). Contrast-modulation was driven by a sinusoidal function (Fig. 1). Following the beginning of the stimulation sequence (background), the stimulus reached its full contrast value after half a cycle (e.g., 250 ms/2 = 125 ms for F = 4 Hz), then decreased back to zero contrast during the remainder of the cycle. A sinusoidal contrast modulation was used because it constrains the input to be at a single temporal frequency and thus the response frequency content of the response spectrum can be better interpreted in terms of harmonics of this single input frequency (Regan, 1972; Van Der Tweel & Lunel, 1965; Victor & Zemon, 1985).

To minimize effects of repetition due to low-level visual cues, the face stimulus changed size at every cycle (random size between 82 and 112% of base face size) in both EEG conditions (Fig. 2a). Four stimulation rates were used (5.88, 7.14, 8.33, 9.09 Hz), corresponding to cycle durations of 100.00, 125.00, 142.86, 150.00 ms, respectively (Fig. 1b). Participants were tested in four separate sessions over four different days. Each recording session consisted of two 84-s stimulation trials for each frequency and condition, for a total of 28 trials. As in our previous studies (Rossion & Boremanse, 2011; Rossion et al., 2012), a long stimulation duration was used, for two reasons. First, with a long stimulation window, one can apply the Fourier transform to a long recording window, so that the frequency resolution of the spectrum is very high. It means that all of the response of interest, and thus all the potential difference between conditions, can be concentrated in a discrete frequency band around the stimulation frequency. This frequency band occupies a very small fraction of the total EEG bandwidth. In contrast, biological noise is distributed throughout the EEG spectrum, resulting in a signal-to-noise ratio (SNR) in the bandwidth of interest between conditions, can be concentrated in a discrete frequency band around the driving frequency. Averaged SNR data for every condition were made separately for each individual participant, and grand-averaged only for display of topographical maps. Regions of interest (ROIs) were selected for statistical analysis based on the maxima of the topographical maps, which are the response identical to our previous face processing study (Rossion et al., 2011; the same channel configuration (Rossion & Boremanse, 2011; Rossion et al., 2012). The ROIs included right/left occipito-temporal cortex (PO9/PO7, PO10h/PO9h, PO10h/ PO9h, PO6/PO5, PO8/PO7, PO9h/PO5h/PO5h, P10/P9, P8/P7, P6/P5) and parieto-occipital medial channels (P2, POZ, P10, POZ), and SNR values at individual channels were averaged over these channels for statistical comparisons.

The recording and analyses

EEG was recorded from 128 Ag/AgCl electrodes mounted in an electrode cap (Wavestad, ANT; for a 2D mapping of electrode labels and positions, see http://www.ant-neuro.com/products/caps/waveguard/layouts/128/). Electrode positions included the standard 10–20 system locations and additional intermediate sites. Vertical and horizontal eye movements were monitored using four additional electrodes placed on the outer canthus of each eye and on the inferior and superior areas of the right orbit. During EEG recording, all electrodes were referenced to a centro-frontal channel (AFz), and electrode impedances were kept below 10 kΩ.

EEG was digitized at a 1000 Hz sampling rate, and a digital anti-aliasing filter of 0.27 Hz was applied at recording (at 1000 Hz, the usable bandwidth is 0 to 270 Hz). All EEG analyses were carried out using Analyzer 2 (Brain Products, Germany); Letswave (http://nocions.webnode.com/letswave) (Mouraux & Iannetti, 2008) and Matlab 7.8. For each individual participant, each EEG record (N =224, including 2 s of data before and after stimulus presentation, was bandpass-filtered between 0.1 and 100 Hz (Butterworth filter, slope = 24 dB/oct). Next, electrodes with artifacts other than eyeblinks were pooled across neighboring channels using linear interpolation. Finally, all channels were re-referenced to a common average reference (without M1/M2 channels). For each 84-s trial, the first 17 s of EEG recordings was removed. This duration corresponds to the beginning of the stimulation (15 s, i.e., baseline) in which the exact same stimulation was presented to the two conditions (identical face repeated), plus two additional seconds of recording. The two additional seconds were excised to avoid including any potential transient ERP component that could have been elicited by the sudden change in global identity – i.e., the onset of the first ‘different F’ stimulus. Roughly 50 s of stimulation from the 18th second (i.e., t = 18 to t = 67 s) were considered for analysis. The rest of the sequence was not used because eyeblinks were more frequent toward the end of recording. The time-window of 50 s was selected for its high and convenient spectral resolution of 1/50 ~ 0.02 Hz (e.g., (Rossion et al., 2012; Srinivasan et al., 2006)). However, to avoid spectral leakage and to ensure a fair comparison across frequency rates, the analysis windows were slightly adapted so that their duration corresponded to an integer number of cycles for every frequency (e.g., 50.008 s for 5.88 Hz and 50.188 s for 3.03 Hz). A Discrete Fourier Transform (Matlab’s DFT) was applied to the resulting individual windows, and EEG amplitude extracted at a resolution of roughly 1/50 ~ 0.02 Hz. SNR was computed at each channel for all frequency bins between 0 and 100 Hz as the ratio of the amplitude at the frequency of interest to the average amplitude of the 20 neighboring bins, skipping only the closest neighboring bin on each side (Rossion & Boremanse, 2011). The results are reported as SNR values, but were qualitatively similar in terms of amplitude at the driving frequency. Average of SNR data for each condition were made separately for each individual participant, and grand-averaged only for display of topographical maps. Regions of interest (ROIs) were selected for statistical analysis based on the maxima of the topographical maps, which are the response identical to our previous face processing study (Rossion et al., 2011; the same channel configuration (Rossion & Boremanse, 2011; Rossion et al., 2012). The ROIs included right/left occipito-temporal cortex (PO10/PO9, PO010h/PO09h, PO10h/ PO9h, PO6/PO5, PO8/PO7, PO9h/PO5h/PO5h, P10/P9, P8/P7, P6/P5) and parieto-occipital medial channels (P2, POZ, P10, POZ) and SNR values at individual channels were averaged over these channels for statistical comparisons.
3. Results

At every stimulation frequency and in both conditions, there were large EEG responses confined to single frequency bins (0.02 Hz) located at the fundamental frequency of stimulation (1F) and its harmonics (2F, 3F,…) (Fig. 2). These responses indicate that the brain synchronized precisely with the rate of visual stimulation, leading to clear ssVEPs (Regan, 1966, 1989). In both conditions, SNR at the fundamental frequency (first harmonic) was by far the largest, with the exception of the second harmonic (2F) response at 3.03 Hz and the third harmonic (3F) response at 2 Hz (i.e., about 6 Hz in both cases).

The fundamental frequency responses were observed mainly at posterior electrode sites, with peaks at either a medial occipito-parietal or a right occipito-temporal location (Figs. 3 and 4). Considering all electrode sites, average SNR values ranged between 4 and 11 (i.e., a 4 to 11 times higher response at the stimulation frequency than at neighboring frequencies). In both conditions, SNR values were the lowest at either very low (1.0 Hz) or high (16.66 Hz) frequency rates.

The difference between the two conditions (different faces—identical face) was restricted to occipito-temporal sites and stimulation frequencies from 4.0 to 8.33 Hz, with a clear right hemispheric dominance (Fig. 5). The maximal difference was at 5.88 Hz for three observers, and at 5.0 Hz for one observer (Fig. 6; see also Table 1 in supplementary material). Despite substantial interindividual variability in SNR, the frequency tuning functions were remarkably similar across individuals (Fig. 6).

An ANOVA with frequency (14 levels) and individual face repetition (same or different faces) as within subjects (4 participants) variables on the SNR of the right occipito-temporal region, averaged the 4 recording sessions per subject showed significant effects of individual face repetition ($F(1, 3) = 17.58$, $p = 0.025$, partial eta$^2 = 0.854$) and frequency ($F(13, 39) = 5.45$, $p < 0.0001$, partial eta$^2 = 0.645$). The interaction between the two factors was significant ($F(13, 39) = 8.80$, $p < 0.0001$, partial eta$^2 = 0.746$), reflecting the higher SNR for different faces than for same faces in the 4.0–8.33 Hz range only (Fig. 6; Tables 1 and 2). Similarly, in the left homologous region there were significant effects of individual face...

![Fig. 2. SNR EEG spectra illustrated for 3 of the 14 frequencies of stimulation: 1.00 Hz (left), 5.88 Hz (right) and 12.50 Hz (below). SNR spectra were extracted from an occipito-temporal region of interest of 9 channels based on topographical maps (Figs. 3 and 4), and grand-averaged over all recording sessions and observers. Clear responses were observed at the fundamental frequency of stimulation and its harmonics. Note the numerous harmonic responses at 1.0 Hz, and the particularly large response observed at 5.88 Hz, with a much higher SNR when different faces are presented as compared to the presentation of the exact same face. Note that the three panels have different y-axis ranges, and that the small depression at the nearby frequencies of the frequency of interest is due to SNR computation across the whole spectra (each frequency bin divided by averaged amplitude of 20 neighbouring bins).](image-url)
repetition \( (F(1, 3) = 44.81, p = 0.007, \text{partial } \eta^2 = .937) \) and frequency \( (F(13, 39) = 2.17, p = 0.031, \text{partial } \eta^2 = .420) \) and a significant interaction \( (F(13, 39) = 3.86, p = 0.001, \text{partial } \eta^2 = .562) \). In the left hemisphere, the SNR was higher for different than for same faces at 4.0–5.88 Hz.

For same faces only, there was no main effect of frequency \( (F(13, 39) = 1.54, p = 0.15, \text{partial } \eta^2 = .399) \) or hemisphere \( (F(1, 3) = 1.63, p = 0.29, \text{partial } \eta^2 = .348) \) and no interaction \( (F(13, 39) = 1.58, p = 0.13, \text{partial } \eta^2 = .345) \). For different faces, there was a effect of frequency \( (F(13, 39) = 7.81, p < 0.0001, \text{partial } \eta^2 = .721) \), no effect of
Fig. 5. Topographical maps (back of the head) of the differential SNR (different faces—identical face) at each frequency of stimulation. Negative values were minimal and set to zero in the display.

Fig. 6. Frequency-tuning functions for each observer (S1 to S4), for each condition and their difference at right occipito-temporal electrode sites (mean and standard errors computed across subjects and sessions). The difference peaks at 5.88 Hz for observers S1, S2, S3, and at 5.0 Hz for S4.
Table 1
SNR values for each observer and each frequency of stimulation at right occipito-temporal cortex ROI. The Z-score is computed as the SNR, on the difference between conditions by considering the amplitude at the frequency bin of interest (i.e., the signal) and the mean and standard deviation of the 20 bins surrounding the frequency of interest (i.e., the noise). In bold, the highest differential SNR values.

<table>
<thead>
<tr>
<th>Stimulation frequency (Hz)</th>
<th>1.00</th>
<th>2.00</th>
<th>3.03</th>
<th>4.00</th>
<th>5.00</th>
<th>5.88</th>
<th>7.14</th>
<th>8.33</th>
<th>9.09</th>
<th>10.0</th>
<th>11.11</th>
<th>12.50</th>
<th>14.28</th>
<th>16.66</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1 Different</td>
<td>6.74</td>
<td>8.20</td>
<td>4.68</td>
<td>6.24</td>
<td>7.98</td>
<td>12.11</td>
<td>8.94</td>
<td>7.75</td>
<td>7.22</td>
<td>6.65</td>
<td>6.14</td>
<td>4.12</td>
<td>3.26</td>
<td>2.10</td>
</tr>
<tr>
<td>Identical</td>
<td>8.12</td>
<td>9.99</td>
<td>4.05</td>
<td>5.12</td>
<td>3.84</td>
<td>4.83</td>
<td>4.92</td>
<td>4.51</td>
<td>5.51</td>
<td>3.72</td>
<td>3.26</td>
<td>3.50</td>
<td>3.56</td>
<td>3.10</td>
</tr>
<tr>
<td>Subtraction</td>
<td>-1.38</td>
<td>-1.79</td>
<td>0.63</td>
<td>1.11</td>
<td>4.13</td>
<td><strong>7.28</strong></td>
<td>4.02</td>
<td>3.23</td>
<td>1.71</td>
<td>2.93</td>
<td>2.88</td>
<td>0.62</td>
<td>-0.30</td>
<td>-1.00</td>
</tr>
<tr>
<td>Z scores</td>
<td>-2.38</td>
<td>-5.13</td>
<td>2.24</td>
<td>3.79</td>
<td>12.71</td>
<td><strong>31.24</strong></td>
<td>17.38</td>
<td>11.99</td>
<td>8.57</td>
<td>8.95</td>
<td>7.34</td>
<td>3.40</td>
<td>-0.03</td>
<td>-3.20</td>
</tr>
<tr>
<td>S2 Different</td>
<td>5.59</td>
<td>5.07</td>
<td>2.56</td>
<td>4.15</td>
<td>5.76</td>
<td>6.57</td>
<td>3.46</td>
<td>1.54</td>
<td>1.91</td>
<td>1.63</td>
<td>2.61</td>
<td>1.92</td>
<td>1.78</td>
<td>1.03</td>
</tr>
<tr>
<td>Identical</td>
<td>4.51</td>
<td>6.67</td>
<td>4.07</td>
<td>3.04</td>
<td>2.67</td>
<td>2.22</td>
<td>2.05</td>
<td>1.34</td>
<td>1.62</td>
<td>1.11</td>
<td>1.63</td>
<td>1.74</td>
<td>1.93</td>
<td>1.39</td>
</tr>
<tr>
<td>Subtraction</td>
<td>1.09</td>
<td>-1.81</td>
<td>-1.50</td>
<td>1.10</td>
<td>3.10</td>
<td><strong>4.35</strong></td>
<td>1.41</td>
<td>0.20</td>
<td>0.28</td>
<td>0.52</td>
<td>0.99</td>
<td>0.19</td>
<td>-0.15</td>
<td>-0.36</td>
</tr>
<tr>
<td>Z scores</td>
<td>0.81</td>
<td>-11.11</td>
<td>-3.55</td>
<td>2.85</td>
<td>10.46</td>
<td><strong>12.47</strong></td>
<td>4.27</td>
<td>0.09</td>
<td>0.75</td>
<td>1.77</td>
<td>4.58</td>
<td>-0.45</td>
<td>-1.11</td>
<td>-1.38</td>
</tr>
<tr>
<td>S3 Different</td>
<td>1.43</td>
<td>2.09</td>
<td>1.93</td>
<td>5.79</td>
<td>7.80</td>
<td>8.83</td>
<td>6.14</td>
<td>4.36</td>
<td>3.67</td>
<td>2.42</td>
<td>4.19</td>
<td>3.02</td>
<td>2.96</td>
<td>3.18</td>
</tr>
<tr>
<td>Identical</td>
<td>1.44</td>
<td>2.06</td>
<td>2.76</td>
<td>3.16</td>
<td>4.44</td>
<td>4.90</td>
<td>3.54</td>
<td>2.90</td>
<td>3.21</td>
<td>2.41</td>
<td>4.23</td>
<td>3.87</td>
<td>3.20</td>
<td>2.59</td>
</tr>
<tr>
<td>Subtraction</td>
<td>-0.01</td>
<td>0.03</td>
<td>-0.82</td>
<td>2.63</td>
<td>3.36</td>
<td><strong>3.93</strong></td>
<td>2.60</td>
<td>1.46</td>
<td>0.46</td>
<td>0.01</td>
<td>-0.04</td>
<td>-0.85</td>
<td>-0.23</td>
<td>0.59</td>
</tr>
<tr>
<td>Z scores</td>
<td>-1.91</td>
<td>-1.28</td>
<td>-3.40</td>
<td>6.63</td>
<td>14.80</td>
<td><strong>13.12</strong></td>
<td>7.06</td>
<td>2.82</td>
<td>1.80</td>
<td>0.64</td>
<td>-1.38</td>
<td>-2.31</td>
<td>-1.12</td>
<td>4.58</td>
</tr>
<tr>
<td>S4 Different</td>
<td>4.05</td>
<td>2.80</td>
<td>4.89</td>
<td>6.38</td>
<td>9.47</td>
<td>8.60</td>
<td>6.72</td>
<td>4.70</td>
<td>3.81</td>
<td>3.48</td>
<td>2.68</td>
<td>2.45</td>
<td>4.19</td>
<td>4.49</td>
</tr>
<tr>
<td>Identical</td>
<td>4.88</td>
<td>4.52</td>
<td>3.65</td>
<td>4.42</td>
<td>4.45</td>
<td>5.21</td>
<td>4.90</td>
<td>2.57</td>
<td>2.47</td>
<td>2.09</td>
<td>1.78</td>
<td>2.85</td>
<td>4.46</td>
<td>2.15</td>
</tr>
<tr>
<td>Subtraction</td>
<td>-0.83</td>
<td>-1.71</td>
<td>1.24</td>
<td>1.96</td>
<td>5.03</td>
<td>3.39</td>
<td>1.82</td>
<td>2.14</td>
<td>1.35</td>
<td>1.39</td>
<td>0.90</td>
<td>-0.40</td>
<td>-0.26</td>
<td>2.34</td>
</tr>
<tr>
<td>Z scores</td>
<td>-4.94</td>
<td>-7.19</td>
<td>5.03</td>
<td>8.20</td>
<td><strong>21.32</strong></td>
<td>12.24</td>
<td>7.50</td>
<td>9.11</td>
<td>2.97</td>
<td>4.31</td>
<td>5.27</td>
<td>-0.55</td>
<td>-1.24</td>
<td>8.55</td>
</tr>
</tbody>
</table>

Significant differences (p < 0.05) are highlighted in bold.

hemisphere (F(1, 3) = 2.57, p = 0.21, partial eta² = .461) and a significant interaction between the two factors (F(13, 39) = 2.93, p = 0.0047, partial eta² = .492). This reflects the higher SNR in the right than in the left hemisphere for frequencies ranging between 5.0 and 8.33 Hz (ps < 0.05). There were no significant differences at other stimulation frequencies (all ps > 0.05; see also Table 2). In addition to the analyses, Table 1 reports significant differences at each frequency rate (Z-scores based on signal and noise at surrounding frequency bins) separately for each individual participant. The only stimulation frequency rates that led to significant effects (Z > 2.33, p < 0.01) in every participant were 4.0, 5.0, 5.88 and 7.14 Hz.

The results were virtually identical when analyses were performed on EEG amplitudes rather than SNR, with a significant difference between conditions at right occipito-temporal sites between 4.0 and 7.14 Hz (all ps < 0.0001; p = 0.07 for 8.33 Hz), and a peak of the tuning function at 5.88 Hz for the same three observers for the ‘different faces’ condition, and for two observers for the difference between conditions.

In summary, stimulation frequencies of 4.0 to 8.33 Hz elicited larger responses when different faces were presented than when the same face was repeated, for all individual participants, with remarkably similar frequency tuning functions. The difference between conditions peaked at around 6 Hz (5.88 Hz) over right occipito-temporal electrode sites. Data analysis in the time-domain showed stable repetitive sinusoidal waveforms and large differences between conditions at 5.88 Hz (Fig. 7). The pattern was similar, though less marked, for other frequencies in the middle range (e.g., 4.0 Hz). By contrast, responses at lower frequency rates were associated with transient components of different shapes, phase and polarity. At higher frequency rates (e.g., 12.5 Hz), the response over occipito-temporal sites was not only of equal magnitude for the two conditions but appeared weaker and desynchronized, with substantial amplitude variations at every cycle.

3.1. Medial occipito-parietal ROI

Although large ssVEP responses were observed at medial occipito-parietal sites, there was only a significant effect of frequency (F(13, 39) = 4.29, p < 0.001, partial eta² = .588), with peaks at 5.88 and 12.5 Hz, but no effect of individual face repetition (F(1, 3) = 0.09, partial eta² = .003) and no interaction between the two factors (F(13, 39) = 1.021, p = 0.45, partial eta² = .254) at these sites (Fig. 8).

3.2. Harmonics

For both conditions, SNR at the first harmonic was by far the largest, with the exception of the second harmonic (2F) response at 3.0 Hz and the third harmonic (3F) response at 2.0 Hz (i.e., 6.0 Hz in both cases) (Table 3). However, there were no differences between conditions at these frequency rates. Differences between conditions were restricted to midrange values (4.0–8.33 Hz) at the first harmonic. The only harmonics for which the differences came close to significance were the second harmonic for the 4.0 Hz stimulus (8.0 Hz response, t = 1.18; p = 0.16) and the third harmonic for the 3.0 Hz stimulation (9.09 Hz, t = 1.99; p = 0.07) (Table 3 in supplementary material).
3.3. Complementary experiment: Varying the number of individual faces

The data reported indicate that at high frequency rates (9.09 Hz onward) presenting different or identical faces leads to identical ssVEP responses. One possibility is that this absence of effect is due to the individual faces being repeated too often at such high stimulation rates in the different faces condition. To rule out this alternative explanation, a complementary experiment was performed with two stimulation frequency rates (5.88 and 12.50 Hz) in five conditions: the same face repeated, or increasing numbers of different faces in the sequences (N = 10, 20, 50, or 100; two repetitions per condition). The data were analyzed for each frequency separately, using an ANOVA with individual face repetition (5 levels) as within subjects (8 participants) variable. At 5.88 Hz stimulation frequency, there was a highly significant effect of individual face repetition (F(4, 24) = 5.3, p < 0.003, partial eta² = .498), with a larger SNR for all ‘different faces’ conditions than when the same face was repeated (all ps < 0.05, Fig. 9). When the ‘identical face’ condition was removed from the analysis, there was no difference between conditions (F(3, 18) = 2.9, p = 0.065, NS, partial eta² = .323). Thus, the magnitude of the difference at 5.88 Hz, which had the same right occipito-temporal cortex topography in all conditions (Fig. 9), did not differ according to the number of different faces used in the experiment. Importantly, there was no effect of individual face repetition at 12.5 Hz (F(4, 24) = 1.96, p = 0.13; NS, partial eta² = .271) and no increase of the effect with an increasing number of different faces (Fig. 9). Overall, this complementary experiment indicates that increasing the number of different faces from 10 to 100 does not influence SNR in that condition, so that this factor cannot explain why differences between ‘different’ and ‘same’ faces disappear at frequency rates above 8.33 Hz (i.e., 120 ms cycle duration) in the main experiment.

4. Discussion

Adaptation to individual faces – as measured by suppression of neural responses to face repetition – is tuned to a relatively narrow range of temporal frequencies ( > 3.03–< 9.09 Hz) centered on approximately 6 Hz. At all suitable frequency ranges, this effect is observed over the right occipito-temporal cortex, a distinct signature of face-specific perceptual processes (e.g., Bentin et al., 1996; Sergent, Ohta, & MacDonald, 1992; Rossion & Jacques, 2008).

Since the specific ~4–8 Hz range of the effect emerges from a comparison of two conditions at equal stimulation frequencies, this result cannot be due to the lower noise level in the theta-band region as compared to the lower (delta) band and higher (alpha) bands of the EEG (Klimesch, 1999; Nunez & Srinivasan, 2006). Moreover, the shape of the frequency-tuning function observed here is unlikely to be due to attentional or arousal processes, which are known to effect ssVEP amplitude (e.g., Morgan, Hansen, & Hillyard, 1996; Muller et al., 2006) and to interact with stimulation frequency (Ding, Sperling, & Srinivasan, 2006). Here, attention was kept constant across frequencies by the orthogonal fixation task, performed at the same level across frequencies. Moreover, an increase of attention would decrease alpha-band amplitude (8–12 Hz: alpha desynchronization; see Klimesch, 1999, 2012) rather than increase specifically.
Finally, a general increase of attention is unlikely to follow a Gaussian-like frequency-tuning curve centered at $C^2_{46}$ Hz, and affect EEG amplitude at specifically the tiny frequency bins of stimulation and at right occipito-temporal sites only.

**4.1. High stimulus presentation rates**

At stimulation rates above 10 Hz, the response focuses on medial occipito-parietal sites in both conditions. This topography,
interrupts or competes (Keysers & Perrett, 2002; Keysers et al., 2001) with its processing. This is consistent with studies showing in higher visual areas, such as the monkey superior temporal sulcus (STS), that a delay of at least 100 ms between a first (target) face and a second (mask) face is necessary to observe two full responses (Gruss, Wieser, Schweinberger, & Keil, 2012; Martens & Gruber, 2012), are used. In contrast, the response at occipito-temporal sites at the stimulus frequency is almost absent above 10 Hz stimulation rates for both face conditions. This suggests that the temporal processing limit of high-level visual cortical areas, even for the low-level features of the stimuli, has been exceeded above 10 Hz.

At rates equal to or above 8.33 Hz (cycle of 120 ms), there are no consistent differences between the two conditions. This observation suggests that at rates \( \geq 8.33 \) Hz, each individual face cannot be processed at a sufficiently fine-grained level before the next face interrupts or competes (Keysers & Perrett, 2002; Keysers et al., 2001) with its processing. This is consistent with studies showing in higher visual areas, such as the monkey superior temporal sulcus (STS), that a delay of at least 100 ms between a first (target) face and a second (mask) face is necessary to observe two full responses of face-selective neurons (Keysers & Perrett, 2002; Kovacs, Vogels, & Orban, 1995; Rolls & Tovee, 1995). In humans, it has also been reported that the ERP response to a second face is absent if it is presented 50 ms after the presentation of a first face (Jeffreys, 1996).

The current results, by showing the temporal limits of a face-identity adaptation effect go beyond previous observations by providing information about the temporal bottleneck of individual face discrimination in humans. Our paradigm tests whether it is possible to discriminate neural responses from same vs. different faces as this is the only factor that differentiates the same vs. different face conditions. The differential response we record over some, but not all frequencies must occur after discrimination of the different exemplars has occurred. The interpretability of our adaptation effect as being indicative of face discrimination compares favorably to psychophysical data obtained by backward masking or rapid visual serial presentation (rsvp) sequences. First, because there are no long gaps between trials, it unlikely that the processing of a highly salient stimulus is facilitated (Keysers et al., 2001). Second, because we measure a neural response that occurs at precisely the time-scale of the stimulus (e.g., a 5.88 Hz response for a 5.88 Hz stimulation rate), we can make more direct statements about temporal processing limits than are possible with indirect measures such as psychophysical backward masking (see VanRullen, 2011). That is, even though an observer might be able to perceive differences among individual faces at high frequency rates, the absence of difference between conditions at such high rates indicates that the face processing system cannot synchronize to, i.e., process, every single face that is presented in the sequence.

4.2. The golden range of stimulation frequencies for face discrimination

At 8.33 Hz and lower rates, the ssVEP response increases over the right occipito-temporal cortex in the ‘different faces’ condition, suggesting that the temporal distance between two different consecutive faces (i.e., 120 ms at 8.33 Hz) becomes sufficient to elicit distinct responses to each individual stimulus, with a delay of 170 ms (5.88 Hz frequency rate) leading to the largest response. This stimulus presentation rate is consistent with timing estimates of evidence accumulation processes in the monkey infero-temporal cortex (about 150 ms for fine-grained discrimination, Matsumoto, Okada, Sugase-Miyamoto, Yamane, & Kawano 2005; Sugase et al., 1999). This 170 ms cycle duration appears to be the minimum duration necessary to fully process an individual face – presented in full color and at 100% contrast – in the occipito-temporal cortex. If this is correct, the maximum differential response should also be observed at about 6 Hz of stimulation rate with other approaches such as human fMRI, and perhaps single-cell or multi-unit recordings in monkey face-sensitive areas. Our findings also carry important practical implications, allowing investigators to optimize the stimulation rates for observing the largest repetition suppression effects to faces and other visual forms in the occipito-temporal cortex (Grill-Spector & Malach, 2001).

Given our observations, we would also speculate that, at the behavioral level, human observers should find it difficult to discriminate faces presented at rates above 8.33 Hz. Obviously, they should not perform better between 4 and 8 Hz than at slower rates. However, it may be that these intermediate frequency rates allow better isolation of the contribution of face discrimination processes to performance. If this is the case, a 4–8 Hz rate may reduce inter-trial and inter-individual variability in behavioral performance at individual face discrimination, and may potentially be used to characterize abnormal processes in patients suffering from face recognition difficulties.

Interestingly, a 170 ms cycle length at this frequency also corresponds to the peak of the face-sensitive N170 ERP (Bentin et al., 1996), which is also the earliest time-point at which individual faces are discriminated over the right occipito-temporal cortex (Jacques & Rossion, 2006; Jacques et al., 2007). This observation suggests that face-related ssVEP responses on the scalp may reflect the linear summation of many successive transient N170 face-sensitive
responses, as proposed for the 40 Hz auditory steady-state response (ASSR) ([Calambos, Makeig, & Talmachoff, 1981; Santarelli et al., 1995]) or sVEP responses to low-level stimuli (e.g., checkerboard reversal, (Capilla, Pazo-Alvarez, Darriba, Campo, & Gross, 2011); see (Regan, 1989)) (see also (Heinrich, 2010) for an illustration of the superposition phenomenon in the steady state response). That is, at frequency rates around 6 Hz, at occipito-temporal sites, the components would combine to form a single, stable, composite wave dominated by the N170, with a difference between conditions accumulating at every cycle (see Fig. 7).

Alternatively, stimulating in the theta range (4–8 Hz) with a peak at about 6 Hz may simply be optimal for visual encoding (of complex stimuli), independently of, or in addition to, time constraints for discriminating complex visual stimuli such as faces. Interestingly, resonances at 6 Hz over occipito-temporal regions were observed not only when stimulating at this fundamental frequency, but also at lower frequency rates leading to 6 Hz harmonic responses (i.e., 6F at 1 Hz, 3F at 2 Hz, 2F at 3 Hz). In all these cases, the response at 6 Hz was associated with the largest amplitude of all harmonics, although no differences between conditions were observed on these harmonic responses. Brain oscillations in this theta band, and in particular at frequency rates at which the effects were found in the present study (4–8 Hz) have been associated with memory encoding ([Buzsaki, Llinas, Singer, Berthoz, & Christen, 1994; Klimesch, Doppelmayr, Russegger, & Pachinger, 1996], prediction of visual detection ([Busch & VanRullen, 2010; Liebe, Hoerzer, Logothetis, & Rainer, 2012]), and accumulation of evidence in perceptual decision making (van Vugt, Simen, Nyström, Holmes, & Cohen, 2012). This functional role of theta oscillations thus generally agrees with our study, which has the advantage of stimulating the visual system at a specific periodic rate, allowing the measure of responses in small frequency bands, experimentally defined, and with a high SNR ([Rossion et al., 2012]).

4.3. The limitations of low frequency rates

The absence of repetition suppression effect at low frequency rates (< 3.03 Hz) may be due to a too long ISI between individual faces, which allows release from adaptation. That is, information about a repeated individual face may remain present in the neural response for no more than 250–300 ms, and disperse thereafter. Interestingly, amplitude modulations of the face-related N170 component to individual face repetition have been found mainly at short ISIs between the adapter and the target face (e.g., 100–300 ms in (Caharel et al., 2009; Jacques et al., 2007)), with studies using longer ISIs generally failing to reveal such effects (e.g., (Jemel, Pisani, Calabria, Crommelinck, & Bruyer, 2003; Jemel et al., 2003; Mnatsakanian & Tarkka, 2004; Schweinberger et al., 2002); see (Rossion & Jacques, 2011) for a discussion of this issue).

Alternatively, the absence of effect at low-frequency rates may be because ‘transient’ ERP components are clearly distinguishable, since there is time to return to baseline before the next stimulus arrives. Differences between conditions may thus take place on successive components of different polarities, so that the overall difference may be cancelled out when measured on the scalp. In other words, differences between individual faces observed at the N170 peak and on successive components (e.g., P200, N250, …) until several hundreds of milliseconds following stimulus onset (see e.g., (Caharel, Jang, Blanz, & Rossion, 2009; Jacques et al., 2007; Schweinberger et al., 2002)) could cancel each other out at the scalp because of their out-of-phase contributions and their projections to different frequency components (i.e., harmonics) (e.g., Fig. 7; 2 Hz stimulation).

If the latter view is correct, differential responses between conditions at such low frequency rates should be observed with other approaches such as fMRI, single-cell or multi-unit recordings in face-sensitive areas. Previous fMRI studies support this view, since ISIs between individual faces are usually of a few hundreds of ms in studies that have revealed repetition suppression effects for individual faces (e.g., (Davies-Thompson et al., 2009; Schiltz & Rossion, 2006)).

4.4. Frequency-tuning, face-specificity and generalization

Although the exact neurophysiological mechanism that explains the human brain preference for certain stimulation frequencies is not yet fully understood, our results suggest that the temporal presentation rate of the input should be such that it prevents the temporal dispersion of the neural response (at too low rates) or the interference (at high rates) between the after-discharge to the first stimulus and the onset-response to the following one. The temporal activation cycle of a population of neurons determines its availability to respond to a given stimulus. When the latter is repeated at fixed time intervals it forces the neurons to respond at a certain rate. If the inter-stimulus interval coincides with the temporal activation cycle, a larger number of neurons can respond to the input and synchronize their responses, causing a noticeable increase in the signal registered on the scalp ([Buzsáki, 2006; Nunez & Srinivasan, 2006]). For complex visual patterns such as individual faces, our findings indicate that the stimulus temporal presentation rate that coincides with the temporal activation cycle of the neuronal assembly specialized in processing these particular stimuli is of about 6 Hz.

Note that previous frequency-tuning EEG studies always used the exact same low-level repeated stimulus (e.g., (Regan, 1966; Srinivasan et al., 2006)), corresponding to the ‘identical face’ condition used here. However, here, at occipito-temporal sites, this ‘identical face’ condition was associated to a relatively flat frequency-tuning function and no significant effect of stimulation frequency. One possible explanation for the lack of peak at 5.88 Hz at these electrode sites for the ‘identical face’ condition is that the optimal frequency rates (4.0–8.33 Hz) for occipito-temporal responses to faces were also those at which the largest repetition suppression effects took place in that condition. Thus, characterizing the temporal frequency-tuning functions of higher level processes and visual areas requires using complex visual stimuli, recording over the occipito-temporal cortex, and presenting different exemplars of the same category in order to minimize repetition suppression effects. Faces may be particularly interesting in this respect because different individual faces are coded by different patterns of responses in the same population of face-selective neurons (e.g., (Abbott et al., 1996; Freiwald, Tsao, & Livingstone, 2009; Leopold et al., 2006; Leopold, Rhodes, Muller, & Jeffery, 2005; Young & Yamane, 1992)), this differential pattern of responses to individual faces allowing the release from visual adaptation. And, because it is the same population of neurons overall – with the same timing characteristics – that is targeted by different faces, the face processing system can be entrained at the periodic rate even when different exemplars are presented in succession ([Rossion et al., 2012]).

The decision to use faces as the image category in this periodic stimulation approach to high-level vision was motivated by several additional factors. Faces form a visually homogenous set of familiar stimuli, which are associated with large and well-defined neural responses. Faces are detected faster and more automatically than other stimuli ([Crouzet et al., 2010; Fletcher-Watson, Findlay, Leekam, & Benson, 2008; Hershler, Golan, Bentin, & Hochstein, 2010; Kiani et al., 2005]) and computer scientists have devoted considerable efforts to build systems that automatically detect faces in images (([Kemelmacher-Shlizerman, Basri, & Nadler, 2008; Viola & Jones, 2004; Yang, Kriegman, & Ahuja, 2002])). Most importantly, unlike most other objects, faces need to be processed at the individual (i.e., discriminated) level to ensure adequate social interactions.
the observation of the largest response over the right occipito-temporal cortex, at the same electrode sites where the face-sensitive N170 component and its modulation by individual faces have been found (Bentin et al., 1996; Rossion & Jacques, 2011), is suggestive of responses at least partially from face-selective populations of neurons. However, our approach is not restricted to face identity and could be extended to determine the frequency-tuning function of other face categorization processes (e.g., face detection, facial expression categorization,...) and other classes of natural images.

Acknowledgments

This research was supported by an ERC grant (faceswep 284025) to BR, JR, JLS and GVP are supported by the Belgian National Fund for Scientific Research (FNRS).

References

Caharel, S., Jiang, F., Blanz, V., & Rossion, B. (2009). Recognizing an individual face: From image information to face identity and could be extended to determine the frequency-tuning function of other face categorization processes (e.g., face detection, facial expression categorization,...) and other classes of natural images.


