All-or-none face categorization in the human brain

Talia L. Retter[1,2], Fang Jiang[2], Michael A. Webster[2] & Bruno Rossion[1,3,4]

1 Psychological Sciences Research Institute, Institute of Neuroscience,
University of Louvain, Belgium
2 Department of Psychology, Center for Integrative Neuroscience,
University of Nevada, Reno, USA
3 Université de Lorraine, CNRS, CRAN - UMR 7039, F-54000 Nancy, France
4 Université de Lorraine, CHRU-Nancy, Service de Neurologie, F-54000, France

Corresponding author:
Talia Retter
1664 N Virginia St.
Mail Stop 296
Reno, Nevada 89557
+1 (775) 303-3472
tlretter@nevada.unr.edu

Abstract

Visual categorization is integral for our interaction with the natural environment. In this process, similar selective responses are produced to a class of variable visual inputs. Whether categorization is supported by partial (graded) or absolute (all-or-none) neural responses in high-level human brain regions is largely unknown. We address this issue with a novel frequency-sweep paradigm probing the evolution of face categorization responses between the minimum and optimal stimulus presentation times. In a first experiment, natural images of variable non-face objects were progressively swept from 120 to 3 Hz (8.33 to 333 ms duration) in rapid serial visual presentation sequences. Widely variable face exemplars appeared every 1 s, enabling an implicit frequency-tagged face-categorization electroencephalographic (EEG) response at 1 Hz. Face-categorization activity emerged with stimulus durations as brief as 17 ms (17 – 83 ms across individual participants) but was significant with 33 ms durations at the group level. The face categorization response amplitude increased until 83 ms stimulus duration (12 Hz), implying graded categorization responses. In a second EEG experiment, faces appeared non-periodically throughout such sequences at fixed presentation rates, while participants explicitly categorized faces. A strong correlation between response amplitude and behavioral accuracy across frequency rates suggested that dilution from missed categorizations, rather than a decreased response to each face stimulus, accounted for the graded categorization responses as found in Experiment 1. This was supported by (1) the absence of neural responses to faces that participants failed to categorize explicitly in Experiment 2 and (2) equivalent amplitudes and spatio-temporal signatures of neural responses to behaviorally categorized faces across presentation rates. Overall, these observations provide original evidence that high-level visual categorization of faces, starting at about 100 ms following stimulus onset in the human brain, is variable across observers tested under tight temporal constraints, but occurs in an all-or-none fashion.

Introduction

A fundamental function of nervous systems is to organize the flurry of sensory inputs generated by their rich, dynamic, and ambiguous environment. This organization requires generating distinct responses to different sensory inputs (i.e., discrimination), but also generating a common response to different but related sensory inputs (i.e., generalization). This categorization function plays a fundamental role in responding adaptively to the environment, and is the foundation of many other cognitive functions, allowing us to learn, memorize, act, and communicate through language, gesture, and expression (Smith & Medin, 1981; Edelman, 1987; Murphy, 2002; Goldstone, Kersten & Carvalho, 2018).

An outstanding issue is whether neural categorization of currently experienced stimuli (i.e., perceptual categorization) is graded or all-or-none. That is, does our brain progressively build categorization responses, or does categorization emerge all at once from an accumulation of non-categorical sensory information? This issue is important insofar as perceptual categorization relates to our subjective experience of the stimulus, or perceptual awareness (e.g., del Cul et al., 2007; Quiroga et al., 2008; Fisch et al., 2009; de Gardelle et al., 2011; Bachmann, 2013; Sekar et al., 2013; Navajas et al., 2014; Windey et al., 2015). Therefore, understanding whether neural categorization is all-or-none may constrain psychological, philosophical, and computational constructs of conscious awareness, as well as advance theoretical models of brain function.

Studies that have addressed this issue have generally relied on vision, the dominant modality in humans and other primates, and used neural measures recorded during variable stimulus viewing conditions. On one hand, a number of studies have found that neural activity in high-level visual cortical areas increases in an analog, continuous way with increases in stimulus visibility (Kovács, Vogels & Orban, 1995; Vanni et al, 1996; Grill-Spector et al, 2000; Keysers et al, 2001; Bar et al., 2001; Moutoussis & Zeki, 2002; Horovitz et al., 2004; Bacon-

Macé et al, 2005; Christensen et al., 2006; Tanskanen et al., 2007), supporting a graded or progressive visual categorization process. On the other hand, other studies have related neural activity to explicit behavioral reports and rather concluded in favor of all-or-none brain responses related to perceptual awareness, more like discrete “hits” and “misses” (Del Cul et al., 2007; Quiroga et al., 2008; Fisch et al., 2009; Harris, Wu & Woldorff, 2011; Shafto & Pitts, 2015). The abundance of evidence on both sides, obtained with vastly different paradigms, has been difficult to reconcile, largely hindering rather than inspiring a coherent theoretical framework.

To illuminate this issue, we introduce an approach that goes beyond previous efforts in a number of ways.

First, going beyond previous studies that compared to detection or identification of a limited set of specific stimuli without exemplar generalization, we truly measure neural *categorization*, i.e., responses that are specific to a certain class of stimuli across a wide range of variable exemplars. To do that, we use a large set of widely variable natural images of faces. Faces are used as the visual stimuli to categorize for a number of reasons. From the first minutes of life and throughout early infancy, the face is arguably the most important visual stimulus for human ecology (Johnson, Dziurawiec, Ellis, & Morton, 1991; Bushnell, 2001; Sugden et al., 2014). Faces are complex multidimensional stimuli that are ubiquitous in our visual environment and drive many of our behaviors. Their initial categorization as “a face” (i.e., generic face categorization) unfolds into an extremely rich social categorization of the individual, allowing to categorize people according to their gender, ethnicity, emotional expression, age, attractiveness, and identity (Bruce & Young, 1998). In neurotypical human adults, generic face categorization evokes extensive neural activity along the (ventral) occipital temporal cortex (e.g., Sergent, Ohta & MacDonald, 1992; Allison et al., 1994; Puce et al., 1995; Kanwisher, McDermott & Chun, 1997; Zhen et al., 2015; Jonas et al., 2016; Grill-Spector et al., 2017; Gao et al., 2018). As shown with face Pareidolia, generic face categorization in humans goes well beyond the identification of well-defined objective physical features of real faces (Caharel et al., 2013; Omer et al., 2019), providing an advantage over artificial systems in terms of generalization within natural and degraded views and environments (Scheirer et al., 2014).

Second, we measure neural face categorization by presenting each face exemplar immediately before and after a nonface object, i.e. visually masked (Helmholtz, 1867; Enns & Di Lollo, 2000). Compared to studies that have used a backward- masking approach to probe perceptual awareness (e.g., Del Cul et al., 2007; Quiroga et al., 2008; Fisch et al., 2009; Sekar, Findley & Linas, 2012; Sekar et al., 2013), we use both backward- and forward-masking. Most importantly, our stimuli are presented in a dynamic visual stream, i.e., variable faces embedded in a rapid train of variable nonface object images for about 1 minute (**Figure 1**), similarly to rapid serial visual presentation (RSVP) sequences (Potter & Levy, 1969; Keysers et al., 2001; 2005; Potter et al., 2014). This paradigm, combined with electroencephalographic (EEG) frequency-tagging (Adrian & Matthews, 1934; Regan, 1989; Norcia et al., 2015) simultaneously measures the two key aspects of generic face categorization, namely *discrimination* by measuring selective responses to face stimuli among non-face objects, and *generalization* by measuring the common selective responses to widely variable exemplars of human faces (Rossion et al., 2015). The neural face categorization responses obtained are devoid of low-level confounds (i.e., amplitude spectrum; Rossion et al., 2015; Gao et al., 2018), and associated with robust neural activity in face-selective regions of the ventral occipito-temporal cortex as shown both with human intracerebral recordings (Jonas et al., 2016) and functional magnetic resonance imaging (fMRI, Gao et al., 2018).

Third and finally, to probe the all-or-none or graded nature of face categorization, we perform two successive experiments with the same participants. In a first experiment, we increasingly sweep stimulus duration from 120 Hz stimulation rate (i.e., 8.33 ms duration) to 3 Hz (333 ms duration). This stimulation mode provides EEG responses specifically tagged to each different stimulation frequency (e.g., 120 Hz…3 Hz). However, irrespective of this progressively decreasing visual stimulation frequency, face stimuli appear every second in the rapid train of variable nonface object images, allowing to objectively identify and quantify the face categorization response at 1 Hz and harmonics (2 Hz, 3 Hz, etc.; Rossion & Retter, 2016a). With this quantification, we are able define the onset and maximal amplitude of the neural face categorization response, as well as inter-individual variability of the response, with respect to face presentation duration. Importantly, we measure categorization implicitly, i.e., without requiring explicit face categorization judgments from the observers, a significant difference with previous studies in which (all-or-none) brain responses related to perceptual awareness are necessarily confounded with behavioral reports (e.g., Del Cul et al., 2007; Quiroga et al., 2008; Fisch et al., 2009; Harris, Wu & Woldorff, 2011; Shafto & Pitts, 2015; see Tsuchiya et al., 2015; Koch et al., 2016).

Previously, graded neural response amplitudes have been reported under varying temporal constraints (e..g, Rolls & Tovee, 1994; Kovács, Vogels & Orban, 1995; Vanni et al, 1996; Grill-Spector et al, 2000; Keysers et al, 2001; Bar et al., 2001; Moutoussis & Zeki, 2002; Bacon-Macé et al, 2005). We aim to replicate this finding in Experiment 1. Then, to evaluate whether it can be explained by a ratio of "all" to "none" neural responses, we perform a second experiment in which we ask the same participants to perform an explicit behavioral face-categorization task under the same variable frequency stimulation conditions. We investigate the stability of participants' neural categorization response across frequency rates between Experiment 1 and Experiment 2. To explore all-or-none neural responses, we investigate the neural responses to behaviorally categorized and non-categorized faces from this second experiment and compare the face categorization responses of Experiment 2 to data from

Experiment 1 to test whether full neural responses to single faces are present at high presentation rates.

As we document below, our observations indicate that while general visual responses can be tagged at the highest presentation rate of 120 Hz over the medial occipital cortex, face categorization emerges significantly as early as 60 Hz (17 ms SOA/duration) for a few trials, but reliably at the group level at 30 Hz (i.e., 33 ms duration in a train of nonface objects) to saturate at 12 Hz (83 ms duration) over occipito-temporal regions. However, there is substantial interindividual variability in the presentation rates for face categorization responses (emerging between 17 and 83 ms for different observers). Importantly, these differences are correlated across behavioral and neural measures. Crucially, our findings reveal that the apparent graded increase in the neural amplitudes of face-categorization responses in fact reflects an all-or-none response occurring occasionally to faces at extremely short viewing times and becoming more consistent, but not more evolved, as presentation duration increases. The tight relationship with explicit behavior indicates that this all-or-none face categorization response corresponds to perceptual awareness.

Materials and Methods

Participants

Sixteen adult participants recruited from the Psychology Department at the University of Nevada, Reno, took part in the experiment. All reported normal or corrected-to-normal visual acuity. Their mean age was 29 years (SE = 1.65; range = 22 to 44), ten identified as female and six as male, and 15 were right-handed. The study was conducted in accordance with the Code of Ethics of the World Medical Association (Declaration of Helsinki, 2013), with protocols approved by the University of Nevada Institutional Review Board.

Stimuli and Display

Stimuli were 112 natural face images and 200 natural object images compiled internally (for examples, see **Movie 1;** Quek et al., 2018; Gao et al., 2018; https://face-categorization-lab.webnode.com/resources). Facial identities were selected to represent a broad range of characteristics (e.g., age, sex, expression, head orientation). Both face and object images were edited from photographs taken under variable environmental conditions and perspectives (in lighting, background, orientation, etc.). Each image was cropped to a square with the target face or object centered or off-center, resized to 200 x 200 pixels, and saved in JPEG format.

Stimulus variations in low-level cues (luminance, contrast, spatial frequency content) were preserved rather than being equalized/normalized between faces and objects to preserve naturalness and because the high amount of variability within both face and object categories reduces the diagnosticity of these cues to any one category (Thorpe et al., 1996; Crouzet et al., 2010; Foldiak et al., 2004; Rossion et al., 2015). Hence, while there are natural differences in amplitude spectrum between faces and non-face objects overall in this stimulus set, these differences do not account for EEG face categorization responses over the occipito-temporal cortex (Rossion et al., 2015; Or et al., 2019) or fMRI face categorization responses in high-level visual regions (Gao et al., 2018), as tested with phase-scrambled images. The stimuli were presented on a CRT monitor (NEC AccuSync 120) with a 120 Hz screen refresh rate and a resolution of 800 x 600 pixels. At the viewing distance of 80 cm, stimuli subtended 7.3° by 7.3° degrees of visual angle. The viewing distance was measured at the beginning of each experiment, with the participants instructed to maintain this distance throughout the experiment, and careful control by the experimenters.

Experiment 1 (EEG Frequency-Tagging) Procedure

After the setup of the EEG system, participants were positioned in front of the monitor, with access to a keyboard with which to start stimulation trials and give responses. The only illumination in the room came from the testing and acquisition monitors. Participants were informed that they would see series of natural images presented at a high frequency rate, which would reduce incrementally throughout each trial. They were not given any information about the types of images they would see. Instead, they were instructed to fixate on a centrally presented fixation cross throughout the experiment, and to press on the space bar when the cross briefly changed luminance, from black to white, 12 times per trial.

Each trial consisted of: 1) the cross to establish fixation lasting between 2-4 s; 2) a 63-s sequence, consisting of nine, contiguous 7-s steps of stimulus presentation at 120, 60, 40, 30, 24, 20, 12, 6, and then 3 Hz; and finally, the fixation cross for another 2-4 s. Images were presented continuously with a 100% duty cycle at maximum contrast (Retter et al., 2018). Frequency rates followed integer multiples of the 120 Hz monitor frame rate, i.e. with 1 frame at 120 Hz (8.3 ms stimulus duration), 2 frames at 60 Hz (17 ms), 3 frames at 40 Hz (25 ms), etc. Sequences were presented consistently in a fast-to-slow frequency-sweep design, in order to target the fastest presentation rate at which faces could be categorized, without faces having already been categorized at slower rates in the sequence. Additionally, this frequency-sweep design encouraged consistent attention to each testing sequence. Throughout each trial, image size varied randomly at each presentation between 95-105% of the original in five steps, further increasing the generalization requirement for the face categorization response. **Movie 1** provides an illustration of the display sequence at the six frequencies from the original experiment which can be displayed with a monitor refresh rate of 60 Hz (60, 30, 20, 12, 6, and 3 Hz).

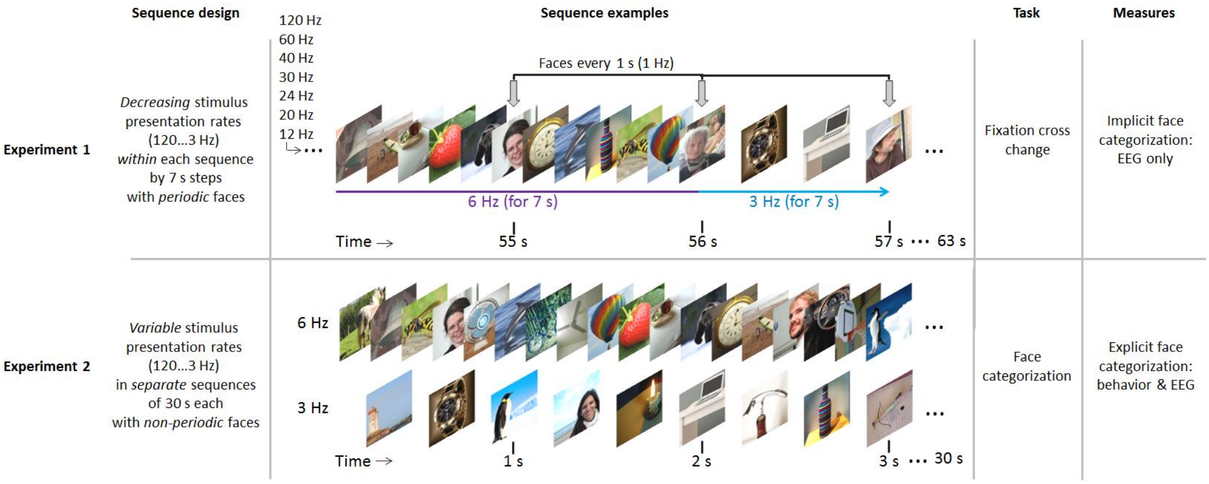


Figure 1. *An overview of the experimental design across Experiments 1 and 2.*

Crucially, regardless of the frequency of image presentation throughout the sequence, faces were embedded consistently at a rate of 1 Hz. Thus, for example, at 120 Hz, faces appeared as every one out of 120 images, while at 6 Hz faces appeared as every one out of 6 images. Providing that the faces are separated by at least 400 ms, the ratio of face to object images does not have an impact on the face-categorization response (Retter & Rossion, 2016a). Besides this constraint of 1 Hz periodicity, the order of the specific face stimuli and the object stimuli was fully randomized for each subject and sequence repetition. Note that although the frequency-sweep approach is novel here, the design to measure generic face (vs. object) categorization with a fixed stimulus presentation rate has been validated in a numerous previous studies, with similar or the same stimulus sets (since Rossion et al., 2015; e.g., Jacques et al., 2016; Retter & Rossion, 2016; Quek et al., 2017; Leleu et al., 2019). Participants binocularly viewed 18 trial repetitions, leading to a total testing time of about 30 minutes.

Movie 1. *An example of a shortened experiment trial, containing the six 7-s frequency steps available for display on a typical 60-Hz monitor (excluding 120, 40, and 24 Hz). Note that despite the incremental decrease in stimulus-presentation frequency, a face appears consistently every 1 s (1 Hz) throughout the sequence. In the icon for the movie, examples of face and non-face stimuli are given: these stimuli were used throughout both Experiments 1 and 2.*

Experiment 2 (Behavior and ERP) Procedure

After Experiment 1, participants were asked two questions: 1) "What did you notice about the images you just saw?"; and 2) "Did you notice any periodicity within the sequences?" In response to the first, open-ended question, participants typically reported seeing a lot of images, the rate of presentation decreasing, and that they noticed faces (14 participants), animals (10 participants), non-living objects (8 participants), and plants/natural scenes (5 participants). Some participants reported one or a few specific images reoccurring: e.g., a dog, elephant, clock, elderly woman, or car. Importantly, in response to the second question, no participants reported noticing any periodicity of faces within the sequences.

Participants were then debriefed in regards to the first experiment, including being informed that faces had appeared every 1 s throughout all of the sequences. They were then given instructions for the second experiment, which instead involved an explicit face detection task. Participants were asked to respond to human faces, which would appear non-periodically, or not at all, within shorter (30 s) sequences, each at a single frequency throughout, by pressing on the key "J" with the index finger of their dominant hand. The faces appeared 5-7 times in five sequences per condition, and 0 times in one sequence per condition. This resulted in a total of about 30 face presentations for each of the nine frequency conditions (M = 30.4; SD = 2.89), for a total duration of about 40 minutes.

EEG Acquisition

EEG acquisition and analysis modeled protocols established in previous studies (e.g., Jacques, Retter & Rossion, 2016; Retter & Rossion, 2016a; Retter et al., 2018). A BioSemi ActiveTwo system (BioSemi B.V., Amsterdam, The Netherlands) with 128 Ag-AgCl Active-electrodes was used to record the EEG. Electrodes were organized in the default BioSemi configuration, centered around nine standard 10/20 locations on the cardinal axes (for exact position coordinates, see http://www.biosemi.com/headcap.htm). For ease of comparison across studies, the default electrode labels were updated to closely match the conventional 10/5 system (for exact relabeling, see Rossion et al., 2015, Figure S2; Oostenveld & Praamstra, 2001). Additionally, vertical and horizontal electrooculogram (EOG) were recorded with four flat-type Active-electrodes: two above and below the right eye and two lateral to the external canthi. The offset of all electrodes was set below 50 mV, relative to the additional common mode sense (CMS) and driven right leg (DRL) electrode loop. The EEG and EOG were digitized at a sampling rate of 2048 Hz and subsequently downsampled offline to 512 Hz.

Analysis

Letswave 5, an open source toolbox (https://www.letswave.org) running over MATLAB R2013b (The MathWorks, USA), was used for the data analysis.

Preprocessing

Data were bandpass filtered from 0.05 to 140 Hz with a fourth-order zero-phase Butterworth filter. Sequences were coarsely segmented from 2 s before to 65 s after sequence onset. To correct for blinks, an independent component analysis was applied for the six participants who blinked more than 0.2 times/s on average throughout the testing sequences (M = 1.20 blinks/s; SD = 0.22; as in, e.g., Retter & Rossion, 2016a). Channels with abrupt deflections of ± 100 µV or above in more than one epoch were linearly interpolated with the neighboring 3-4 channels (M = 0.75 channels; SE = 0.86 channels; range = 0-3 channels, i.e., up to no more than 3% of channels).

Data were re-referenced to the average of all 128 EEG channels. Each 7-s presentation duration step was re-segmented in two ways: 1) beginning from -200 ms before the first face stimulus onset, to capture face-categorization responses; and; 2) beginning exactly at face onset, in order to isolate the stimulus-presentation response to the full 7 s of each presentation duration step.

Regions-of-interest

To determine significance of the face-categorization responses, a ten-channel bilateral occipito-temporal (OT) ROI was defined a priori based on previous studies with the generic face categorization paradigm (e.g., Retter & Rossion, 2016b; Retter et al., 2018; see also Dzhelyova & Rossion, 2014). To decompose the responses further, the average amplitude of the right and left OT regions were examined separately (right: channels P10; P8; PO8; PO10; PO12; and left: P9; P7; PO7; PO9; PO11). The bilateral ROI was verified to capture face-categorization responses post-hoc (see Results: **Figure 1B; Figure 5B**). To provide a region-free assessment of response significance, we also investigated responses across the average of all 128 EEG channels. Finally, to identify and quantify general responses to visual stimulation, we used a medial occipital (MO) region consisting of the 10 following channels: O2; POI2; I2; Iz; OIz; Oz; POOz; O1; POI1; I1), as well as the 128-channel grand average (also verified post-hoc and contrasted to the OT ROI; **Figure 1A; Figure 1C; Supplemental Table 1A; Figure 5D)**.

Frequency Domain Analysis (Experiment 1)

The segmented data were averaged by presentation duration step in the time domain, reducing non-phase locked (i.e., non-stimulus related) activity and thereby increasing the signal-to-noise ratio. A fast Fourier transform (FFT) was applied to represent the data of each channel in the temporal frequency domain, i.e., as a normalized amplitude spectrum (μV), with a range of 0 to 256 Hz and a resolution of 0.14 Hz. To correct for variations in baseline noise level around each frequency of interest (i.e., 1 Hz and its harmonics for face-categorization responses, and the stimulus presentation rate and its harmonics), the average amplitude of the neighboring six frequency bins were averaged (a range of 0.86 Hz) and subtracted from each frequency bin. The harmonics of the face-categorization responses were subsequently summed from the fundamental, 1 Hz, up to 20 Hz (similarly to Retter & Rossion, 2016a; Retter et al., 2018), excluding harmonics that coincided with the stimulus presentation frequency when present. The harmonics of stimulus-presentation responses were summed from the fundamental up to 120 Hz for all presentation duration steps.

To assess significance of the face-categorization and stimulus-presentation responses at each presentation duration condition at the group level, a Z-Score was computed on the average of the bilateral OT ROI channels for face-categorization responses, and the average of the MO ROI channels for the stimulus-presentation response (using the same baseline frequency range as for the baseline-subtraction). Responses were considered to be significant when the Z-Score exceeded 2.32 ($p<.01$, 1-tailed, i.e., predicting the frequency-tagged signal amplitude was greater than the neighboring noise). Note that the frequency-domain amplitude spectrum contains only positive amplitude values beyond 0 Hz. Additionally, the same approach was used at the individual participant level, with a Z-Score threshold of 1.64 ($p<.05$).

To describe the data further, the amplitude and scalp topographies were assessed across presentation conditions. Scalp topographies were visualized in terms of both their original amplitude and their normalized amplitude (McCarthy & Wood, 1985). To assess the extent of right lateralization for face categorization, a lateralization index was computed using the amplitudes of the right and left OT ROIs (R and L, respectively) as follows: $(R-L)/(R+L)*100$

(as in Retter & Rossion, 2016a). Statistical tests on response amplitudes were applied independently for face-categorization and stimulus-presentation responses in the form of repeated measures analysis-of-variance tests (ANOVAs), with factors of *Region* and *Condition*. A Greenhouse-Geisser correction was applied to the degrees of freedom in the event that Mauchly's test of sphericity was significant. To compare across adjacent frequency steps in subsequent post-hoc analyses, paired-sample, two-tailed, t-tests were applied only on the eight adjacent frequency conditions, with a Benjamini-Hochberg correction applied for multiple comparisons. Data were grand-averaged across participants for display purposes.

Time Domain Analysis (Experiments 1 & 2)

Segmented data were filtered more conservatively with a fourth-order, zero-phase Butterworth low-pass filter at 30 Hz. A notch filter was applied to remove the periodic responses to stimulus-presentation for each presentation condition at its fundamental and harmonic frequencies up to 30 Hz. Data were then re-segmented as described previously for face-categorization responses and: either 1) in Experiment1, averaged by condition; or 2) in Experiment 2, given that this experiment had relatively few trials per condition (about 30 on average, containing a mix of categorized and non-categorized faces), we combined responses across a range of presentation rates. Specifically, we used 12 to 60 Hz, in order to include a large number of trials, balanced in terms of the number of categorized (M = 62, SE = 7.1) and non-categorized (M = 60, SE = 5.7) faces across participants (after artifact rejection for muscular artifacts associated with blinks). In a follow-up analysis in Experiment 2, we combined the 12 and 20 Hz conditions and the 24 and 30 Hz conditions, for nine participants who had a minimum of 20 artifact-free face categorizations at both frequency groups; a mean of 33 categorized faces at 12 and 20 Hz (SD = 8.9), and 29 categorized faces at 24 and 30 Hz (SD = 8.5). In all cases, a baseline correction of voltage offset was applied on the 200 ms immediately preceding face stimulus onset.

To test for statistical significance of the response from the baseline reference interval (-200 to 0 ms) at each time point, bootstrapping was performed with a significance criterion of $p<.001$ (two-tailed, to test positive or negative signal deflections) with 10,000 iterations. To reduce the likelihood of false positives, only groups of at least 5 consecutive time bins (i.e., about 10 ms) were reported. In Experiment 1, data were analyzed up to 600 ms post face-stimulus onset; in Experiment 2, face-categorization responses were confounded with behavioral detection (pre-)motor activity, with response times occurring on average at 535 ms over these presentation rates (SE = 15.1 ms). We thus focused the analyses on the first two face-selective components, the P1-face (onset at about 100 ms, 144-ms peak latency) and N1-face (202-ms peak latency), as defined in Experiment 1 (Section 4.2; and reported in previous studies with this paradigm, see e.g., Rossion et al., 2015; Retter & Rossion, 2016a; Quek & Rossion, 2017; Or et al., 2019). To further probe a potential null result, we obtained the Bayes factors corresponding to tests of the samples' means (t-tests; Rouder et al., 2009) with the 'BayesFactor' package (http://bayesfactorpcl.r-forge.r-project.org/) in R. In both experiments, grand-averaging across participants was computed for display.

Additionally, to describe a face-categorization response independent of the stimulus presentation duration, the data of Experiment 1 were averaged across all presentation-duration conditions yielding significant frequency-domain face-categorization responses.

Behavioral Face-Categorization Analysis (Experiment 2)

Face-categorization responses were considered correct when occurring between 0.15 – 2 s after face presentation onset; responses occurring outside this range were considered false positives. This criterion rejected 3.5% of potentially correct face-categorization responses, and generated a response time distribution with a similar mean and median: 509 (SE = 117.5 ms) and 512 ms, respectively. False positives occurred on average at a rate of 0.24 per minute (SE = 0.087); they occurred maximally in a mid-frequency range, from 20 – 40 Hz, although this may be for several reasons (e.g., that participants found the task too difficult to try at the highest frequencies, or that they actually did perceive faces sometimes in the stream of non-face objects (pareidolia); **Supplemental Figure 5**). Note that this generous window for correct responses was chosen in terms of the minimum and maximal expected response times for any participant for any key press (the minimal time between two consecutive face images was 2.6 s); however, application of a more conservative range from 0.2 – 1 s after face presentation only minimally affected the pattern of results across participants or frequencies (e.g., **Supplemental Figure 5**, for the effect on false positives).

The total percent accuracy was defined as the number of correct detection responses minus the number of false positives, relative to the total number of faces presented. This modification of the raw accuracy guarded against the chance that a high rate of correct responses occurred as a byproduct of a high rate of false positives; this was particularly important given the differences in false alarms across presentation rates. Response time (RT) was considered only for correct responses. Statistical differences across stimulus presentation duration conditions were assessed with one-way analysis-of-variance tests (ANOVAs). As for amplitude, post-hoc, paired-sample, two-tailed, t-tests were applied only across the eight adjacent frequencies, with a Benjamini-Hochberg correction applied for multiple comparisons.

Additionally, inverse efficiency (IE) was computed by dividing RT by total percent accuracy (corrected for false positives) for each participant (Townsend & Ashby, 1983). This IE measure was used to rank participants in their face-categorization ability overall, since it combines accuracy and RT performance in a single measure, by using the average of trials from all presentation conditions together. Note that IE is thus used only to rank participants' face-categorization *ability*, while in all other analyses accuracy is used in order to examine *whether or not* faces were categorized, irrespective of the RT. The relationship between behavioral face-detection ability and EEG responses was tested with Pearson correlations between IE and EEG response amplitude. Differences in correlation coefficient strength were calculated according to the asymptotic z-test of Lee and Preacher (2013).

Behavioral Fixation-Detection Analysis (Experiment 1)

Responses to the detection of the fixation cross luminance change were analyzed with the same parameters as for the face-categorization responses. Overall, participants' accuracy was high for this easy task across all frequencies (93.4%; SE = 1.39%), with a mean correct RT of 480.0 ms (SE = 16.27 ms). The rate of false positives was low: 0.15 per minute (SE = 0.038). The accuracy was higher at the highest frequencies (above 96% at 30 Hz and above; relative to 91 – 93% at all lower frequencies), although this could be either because the luminance changes of the cross were easier to see against a more rapid (i.e., less well-processed) image stream, or as an order effect, with the higher frequencies consistently appearing first in the sequences. Overall, the high performance at the fixation cross task supported that participants were maintaining attention to the center of the screen. Moreover, the fixation cross task results differed greatly from those of the face categorization task, emphasizing the differences in these tasks and processes.

Results

To organize the results in relation to the experimental questions tested, the results are divided in five sections. In Section 1, we investigate the minimal and optimal image presentation time for face categorization. To do so, face-categorization responses are characterized across presentation rates from 120 to 3 Hz, both at a neural level from the frequency-tagged EEG responses to periodic faces in Experiment 1, and at a behavioral level from the face-categorization response accuracy to non-periodic faces in Experiment 2. In Section 2, we ask whether responses to behaviorally categorized faces in Experiment 2 produce all-or-none neural responses. The data support all-or-none neural responses, which at first may appear inconsistent with the graded amplitude reported for Experiment 1 in Section 1 across presentation rates. Thus, in Section 3, we explore the relationship between accuracy and EEG amplitude to probe whether all-or-none neural responses could produce the graded amplitudes reported here and in previous studies as a function of stimulus presentation time. While all-or-none neural responses would predict lower amplitudes at high stimulus presentation rates, they would not predict qualitative changes in the responses: thus, a qualitative assessment of face-categorization responses across presentation rates is presented in Section 4. Finally, in Section 5, we test whether all-or-none neural responses might also be used to predict individual differences in generic face categorization.

1. Minimal and Optimal Stimulus Presentation Time for Face Categorization

1.1 Minimal Stimulus Presentation Time for Face Categorization (EEG: Experiment 1; Behavior: Experiment 2)

Neural amplitudes as a function of stimulus presentation time were examined for face-categorization (1 Hz and its harmonics; **Figure 2**) by summing across baseline-subtracted harmonics (Retter & Rossion, 2016a; see Methods) in Experiment 1, recording implicit neural face-categorization responses. Behavioral responses were analyzed in terms of minimal accuracy in Experiment 2, recording explicit behavioral face-categorization responses.

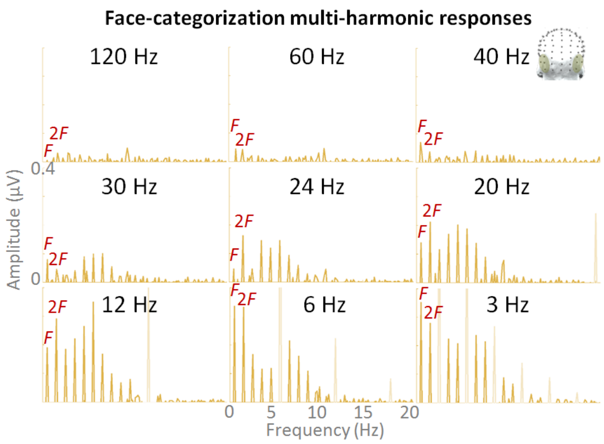


Figure 2*. Face-categorization responses characterized in the frequency domain across conditions (Experiment 1), in the form of baseline-subtracted amplitude spectra. Face-selective responses were tagged at F (= 1 Hz) in every condition, with higher harmonics occurring at 2F (2 Hz), 3F (3 Hz), etc. (only the first two harmonics are labeled above the spectra). The data are shown over the region of the scalp (bilateral occipito-temporal ROI) and frequency range (up to 20 Hz) that were selected to capture face-selective responses.*

Higher harmonic frequencies coinciding with stimulus-presentation harmonic frequencies are drawn in very light yellow, and may surpass the plotted amplitude range. For ease of comparison across conditions, all graphs are plotted with common axes.

At the individual participant level, significant neural OT face-categorization response amplitudes (Z-score criterion: see Methods) emerged between a range of 60 to 24 Hz (17 to 83 ms; **Figure 3A**). On the high end, a single participant had a significant EEG response at 60 Hz (two participants at 40 Hz); on the low end, a single participant had a significant EEG response only up to 12 Hz (three participants only up to 20 Hz). The mean across participants was 29.4 Hz (SE = 2.73 Hz), and the median 30.0 Hz. Individual Z-Score values, as well as scalp topographies and occipito-temporal amplitudes, are exemplified at 30 Hz in **Supplemental Figure 1A**. In terms of behavioral face-categorization responses, only one participant had more correct face categorizations than false alarms at 120 Hz (corresponding to 3.2% accuracy after adjustment), while eight participants reached this level at 60 Hz (ranging from 3.5-12% accuracy), and at least 15 participants at all slower presentation rates (**Figure 3B**). In sum, at the individual participant level, significant categorization responses were first present in both neural and behavioral measures at 17 ms of stimulus presentation duration (60 Hz).

At the group level, while neural face-categorization responses emerged at 40 Hz (25 ms) over OT regions, they did not reach statistical significance at the group level until 30 Hz (33 ms) over the occipito-temporal ROI ($Z = 3.5$, $p = .0002$; **Figure 3A;** also **Figure 4A; Supplemental Table 1A**). Note that the robust statistical significance of neural face-categorization responses, even at the shortest significant stimulus presentation duration, would not have been eliminated had stricter testing criteria been applied (the p-value of .0002 is 50 times lower than the significance criterion of $p = .01$). Below 30 Hz, these responses remained significant ($Z > 2.32$; $p<.01$) at all stimulus presentation rates (all p's < .00001). Behaviorally, face-categorization emerged at 60 Hz at the group level (17 ms; M = 3.03%; SE = 1.26%), t_{15}

= 2.41, p = .015, one-tailed (testing whether the accuracy > zero), d = 0.85 (**Figure 3B;** also **Table 1A; Figure 6A**). At the higher frequency step, 120 Hz, accuracy did not differ from 0% (-0.40%; SE = 0.50%); at the lower frequency step, 40 Hz, accuracy jumped to 22% (SE = 3.2%). The accuracy was 43% at 30 Hz, the frequency rate at which the neural, group-level face categorization response emerged significantly in Experiment 1. The relatively longer presentation duration required for group-level face-categorization response significance measured with EEG (33 ms) than behavior (17 ms) may be accounted for by a very low number of behavioral face categorizations at short durations (vs. non-categorized faces) across participants, and a conservative identification of neural signal *vs.* baseline activity (noise) applied here (see Discussion).

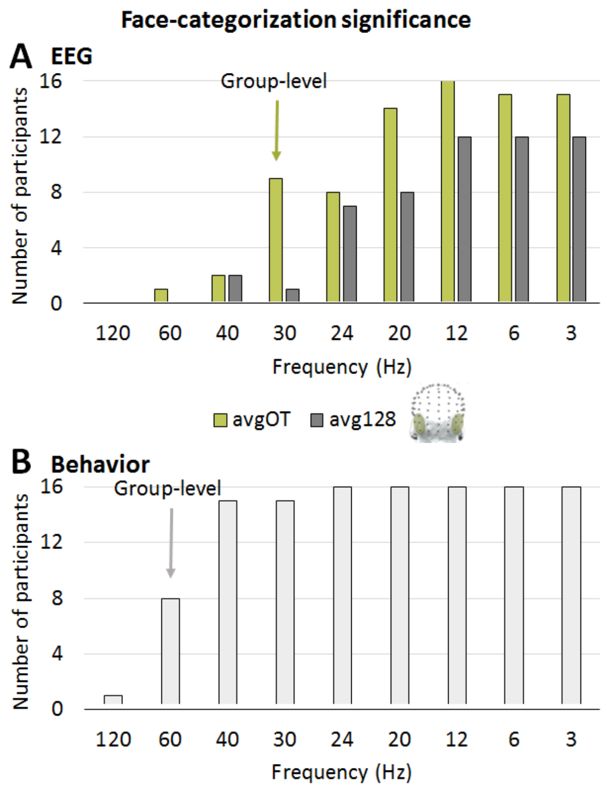


Figure 3*. Face-categorization responses as a function of stimulus presentation rate.* ***A)*** *The number of participants (out of 16 in total; Experiment 1) with significant EEG responses at each stimulus presentation rate (Z > 1.64; p< .05). Group-level significance first emerged at 30 Hz over the occipito-temporal ROI (p = .0002).* ***B)*** *The number of participants with more behavioral categorization hits than false alarms (Experiment 2). Group-level significance first emerged at 60 Hz (p = .015).* ***Key)*** *avgOT = average across the left and right occipito-temporal ROIs; avg128 = average of all 128 EEG channels.*

1.2 Optimal Stimulus Presentation Time for Face Categorization (EEG: Experiment 1; Behavior: Experiment 2)

At a neural group-level, face-categorization response amplitude varied as a function of stimulus presentation duration (Experiment 1; **Figure 4A**; for individual participant data, see **Supplemental Figure 1B**). There was a main effect of *Frequency*, $F_{3.3,49} = 5.94$, p = .010, $\eta_p^2 = 0.86$, with almost no amplitude at 120 and 60 Hz (M ≤ .02 µV), and then a gradually increasing amplitude from 40 to peak at 12 Hz (83 ms; M = 1.50, SE = 0.20), and then slight decrease at 6 and 3 Hz. (This pattern in amplitude was further verified at the average of all 128 channels, as well as at the targeted occipito-temporal ROIs; however, the decrease in amplitude from 12 Hz to 6 and 3 Hz was not present across all 128 channels or systematically at the LOT; **Figure 4A)**. There was also a significant main effect of *Region*, $F_{1,15} = 19.5$, p < .001, $\eta_p^2 = 0.57$, reflecting the larger amplitude for the OT (M = 0.94, SE = 0.12) than MO (M = 0.56, SE = 0.06) ROI in selective responses to faces (as expected for face-categorization responses; please find the scalp topographies in **Figure 4B**; see also **Figure 4C,** contrasting directly the OT and MO regions). The interaction between *Frequency* and *Region*, $F_{3.6,54} = 10.2$, p = .002, $\eta_p^2 = 0.91$, was significant: there was no difference in amplitude across regions from 120-40 Hz (all differences < 0.02 µV), but there appeared to be a separation at 30 Hz and all lower frequencies (all differences > 0.32 µV). This was supported by a statistically significant reduction in amplitude over the OT region from 30 to 40 Hz, $t_{15} = 3.48$, p = .003, d = .89, and no significant differences between 40 and 60 Hz (p = .031; critical value = .025) or 60 and 120 Hz (p = .50). Moreover, post-hoc tests revealed a marginally significant decrease in amplitude over the OT region from 12 to 6 Hz, $t_{15} = -2.36$, p = .032 (critical value = .031), d = .31, and a significant decrease from 12 to 20 Hz, $t_{15} = 2.92$, p = .011, d = .56 . Thus, the presentation duration that produced the optimal, i.e., largest, face-categorization response amplitude over the OT region was 83 ms (12 Hz).

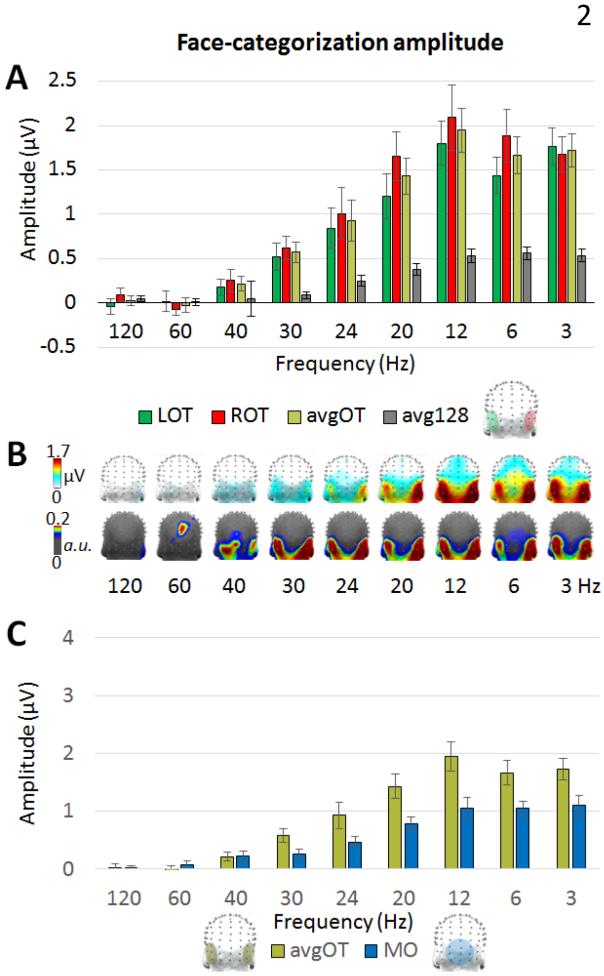


Figure 4. ***A****) Neural face-categorization responses (baseline-subtracted, summed-harmonic amplitudes) as a function of stimulus-presentation rate (Experiment 1).* ***B****) Scalp topographies (top row: amplitudes; bottom row: normalized amplitudes).* ***C****) The bilateral occipito-temporal (OT) and the medial occipital (MO) ROIs for face-categorization responses.* ***Key)*** *LOT: left occipito-temporal ROI; ROT: right occipito-temporal ROI; avgOT = average of these left and right occipito-temporal ROIs; avg128 = average of all 128 EEG channels.*

At the behavioral group-level, face-categorization responses were also affected by stimulus presentation rate (Experiment 2). Overall, the differences in behavioral accuracy, adjusted for false-positives (se Methods) across frequency rates were highly significant, $F_{2.2,33} = 217$, $p<.001$, $\eta_p^2 = 0.94$. Accuracy appeared to be at 0% at 120 Hz, and to increase steadily up to up to 68% at 20 Hz (50 ms), after which it neared ceiling from 12 to 3 Hz (83 to 333 ms; all accuracies above 95.6%) (**Table 1A**; for individual participant data, see **Supplemental Figure 1C**). Post-hoc tests revealed that significant differences were not present between 3 and 6 Hz and 6 and 12 Hz (both p's > .28), but were present across all other adjacent frequency steps (all p's ≤ .02). Therefore, accuracy was optimal at presentation durations from 83 – 333 ms (12 to 3 Hz).

Frequency (Hz)	120	60	40	30	24	20	12	6	3
A. Accuracy (%)	-0.4 (0.49)	3.0 (1.26)	21.7 (3.17)	42.9 (5.59)	57.9 (5.20)	68.0 (4.32)	95.6 (2.12)	97.8 (0.86)	98.0 (1.01)
B. RT (ms)	1435 (145)	650 (52.7)	557 (19.2)	529 (18.9)	505 (11.0)	503 (15.2)	498 (13.4)	495 (15.3)	499 (15.2)
C. IE	84.7 (260)	74.5 (20.8)	32.0 (4.85)	14.2 (1.91)	10.1 (1.16)	8.11 (0.81)	4.29 (0.26)	5.07 (0.18)	5.11 (0.18)

Table 1*. Behavioral face-detection results (Experiment 2): percent accuracy (****A****), response time (RT;* ***B****), and inverse efficiency (IE;* ***C****). One standard error for each measurement is given in parentheses.*

In considering behavioral optimality, response time (RT) was also taken into account. Response time (RT) showed less variation across presentation rates (**Table 1B**). Significant differences across presentation durations were driven only by slower responses at the shortest presentation durations. At 120 Hz, the response was greatly delayed, although note that only two participants had any (one or two) responses at this frequency, such that it could not be statistically compared. At the next shortest presentation rate, 60 Hz, 10 participants had some correct categorization responses, contributing towards significant RT differences across frequency conditions from 3 to 60 Hz, $F_{1.5,12} = 4.70$, p = .040, $\eta_p^2 = 0.37$. These differences persisted when removing 60 Hz, $F_{6,78} = 6.84$, p <.001, $\eta_p^2 = 0.35$, in this case driven by slower responses at 40 Hz. However, from 42 to 333 ms (24 to 3 Hz), mean RTs ranged from 495-505 ms, and there were no significant RT differences across conditions, $F_{4,60} = 0.37$, p = .83, $\eta_p^2 = 0.02$. Although RT was thus not highly sensitive to stimulus presentation duration, being optimal across a wide range from 24 to 3 Hz, RT was taken into account in considering individuals' overall behavioral performance in the measurement of inverse efficiency (IE, i.e., RT/accuracy: see Section 5.1; **Table 1C**).

Overall, given such stable behavioral responses times above 42 ms (24 Hz) stimulus-presentation durations, and near-ceiling behavioral accuracy above 83 ms (12 Hz), optimal behavioral face-categorization responses occurred from 83 to 333 ms (12 to 3 Hz). In agreement, the optimal neural responses fell within this presentation duration range, at 83 ms (12 Hz).

1.3 Comparison with Stimulus-Presentation Responses (Experiment 1)

Neural stimulus-presentation responses, common to all images, were also affected by presentation rate in Experiment 1 (**Supplemental Figure 2** for multi-harmonic responses**; Supplemental Figure 3** for summed-harmonic amplitudes and scalp topographies; see also **Supplemental Table 1B** for Z-scores**).** These responses were always above zero, even at the highest 120 Hz frequency (see Herrmann, 2001 for previous evidence of up to 100 Hz frequency-tagged responses). They appeared to increase exponentially as the presentation rate decreased, at least until the lowest 3-Hz frequency tested. Thus, they exhibited differing dynamics than for face categorization (compare **Supplemental Figure 3A** with **Figure 4A**), although also producing a main effect of *Frequency*, $F_{1.5,22} = 21.8$, $p < .001$, $\eta_p^2 = 0.96$. In opposition to the face-categorization responses, stimulus-presentation responses were larger over the MO (M = 0.91, SE = 0.07) rather than OT (M = 0.63, SE = 0.04) ROI, leading to a significant main effect of *Region*, $F_{1,15} = 43.8$, $p < .001$, $\eta_p^2 = 0.75$ (**Supplemental Figure 3C; Supplemental Table 2B**). The interaction between *Frequency* and *Region*, $F_{1.8,26} = 21.8$, $p = .001$, $\eta_p^2 = 0.92$, was significant, as there was no amplitude difference across regions from 120-60 Hz (differences < 0.02 µV), but an increasing difference as the frequency decreased: at 40 Hz up slightly to 0.035 µV, at 30 Hz, 0.11 µV, and at all lower frequencies > .20 µV. Here, the presentation rate producing the largest stimulus-presentation response amplitude was thus

3 Hz. Overall, distinct localization and pattern of face-categorization and stimulus-presentation responses across stimulus-presentation rates shows the specificity of these response types, and supports the subsequent focus on the presentation times for face-categorization responses over the most sensitive OT regions.

2. Are Neural Responses to Behaviorally Categorized *vs.* Non-categorized Faces All-Or-None?

2.1 EEG Responses to Behaviorally Categorized or Non-categorized Faces (Experiment 2)

In order to investigate whether neural responses to behaviorally categorized *vs.* non-categorized faces are all-or-none or graded as stimulus presentation rate changes, we examined non-periodic EEG responses in the time domain from Experiment 2 separately for 1) behaviorally categorized faces and 2) behaviorally non-categorized faces (see Methods).

For behaviorally categorized faces, significant neural responses (event related potentials) were present during the windows of the previously described category-selective deflections P1-face and N1-face (Rossion et al., 2015; Retter & Rossion, 2016a). For non-categorized faces, no responses were apparent, exhibiting no significance (**Figure 5A; 5B**). Examination of the sum of the P1-face and N1-face peak-time deflections over the bilateral OT ROI also showed a significantly larger response for behaviorally categorized over non-categorized faces, $t_{15} = 6.73$, $p < .001$, $d = 2.0$, Bayes factor of 3,162. Moreover, there was a significant response to behaviorally categorized faces ($M = 3.12$ μV, $SE = 0.37$ μV), $t_{15} = 8.47$, $p < .001$, $d = 3.0$, Bayes factor of 38,492. However, there was not a significant response to behaviorally non-categorized faces ($M = 0.38$ μV, $SE = 0.30$ μV), $t_{15} = 1.24$, $p = .12$, $d = 0.44$, Bayes factor of 0.49 (**Figure 5B**).

The scalp topographies of these face-selective components for behaviorally categorized faces were maximal over the OT region; no clear activation was observed across the scalp to faces that were not behaviorally categorized (**Figure 5C**). Moreover, behaviorally non-categorized faces elicited no significant responses even when investigated over a longer time window, up to 600 ms post-stimulus onset (**Figure 5D**).

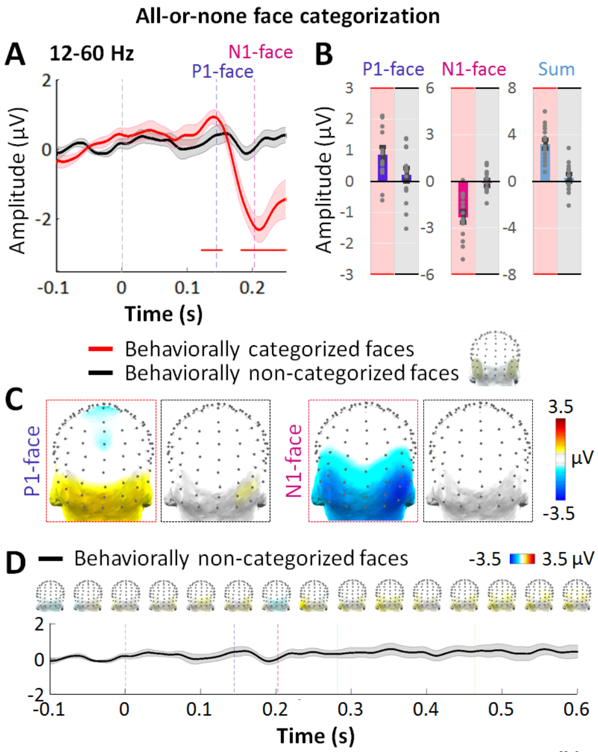


Figure 5. *Time-domain responses (N = 16) across 12-60 Hz to all behaviorally categorized faces (Experiment 2).* ***A)*** *Waveforms over the occipito-temporal ROI, to behaviorally categorized (in red) and non-categorized (in black) faces. Significant time points (p<.001) are indicated below for categorized faces (in red) and non-categorized faces (none). Shading indicates ± one standard error of the mean.* ***B)*** *The amplitude at the selective P1-face and N1-face components, as defined in Experiment 1 (see Figure 6A), as well as their sum, at the occipito-temporal ROI. Group-average data is plotted in bar graphs, with error bars of ± one standard error of the mean, and individual participant data is plotted in gray dots. Summed responses to behaviorally categorized faces (red highlight) are significantly different from zero at the group-level, while those to faces that were not categorized (black highlight) are not significant.* ***C)*** *Posterior topographies at the peak times of the P1-face and N1-face, for behaviorally categorized (red outline) and non-categorized (black outline) faces. Component times, taken from Experiment 1, are indicated in Panel A.* ***D)*** *Responses to non-categorized faces are absent over a longer time window (no significant time points). Topographies above the occipito-temporal waveform are sampled every 50 ms from 100 ms prior to stimulus onset.*

3. All-Or-None Neural Face-Categorization Responses Can Explain Graded Response Amplitudes across Presentation Rates

How can all-or-none neural responses, as reported in the previous section for data collected in Experiment 2, be reconciled with the gradient of response amplitudes reported in Experiment 1? We hypothesize that the gradient of amplitude in Experiment 1 simply reflects the *proportion* of absent ("none") neural responses to non-categorized faces and full ("all") neural responses to categorized faces, in the same way that the behavioral accuracy reflects the proportion of "hits" and "misses". We tested this hypothesis in two ways: first, by examining the overall relationship of amplitude (quantified in the frequency domain; Experiment 1) and accuracy across presentation rates (Experiment 2); and second, by examining the relationship of amplitude (quantified over time-domain components) for behaviorally categorized and non-categorized faces and accuracy (Experiment 2 only).

3.1 The Correlation of Accuracy (Experiment 2) and Amplitude (Experiment 1)

If the gradient amplitudes reported in Experiment 1 are due to missed detections, there should be a linear relationship between amplitude and accuracy. That is, if full categorizations (near-ceiling accuracy) produce a full neural response, then failing to categorize, e.g., 10%, of those faces should reduce the neural amplitude by 10%.

In line with this prediction, the face categorization accuracy (Experiment 2) and EEG amplitudes (Experiment 1) showed similar patterns across stimulus presentation rates (emphasized in **Figure 6A**). Moreover, there was a very strong correlation across all nine presentation rates, $R^2 = 0.96$, $p < .0001$, only slightly stronger in the range where amplitude consistently varied, 60-12 Hz, $R^2 = 0.97$, $p = .0002$ (**Figure 6B**).

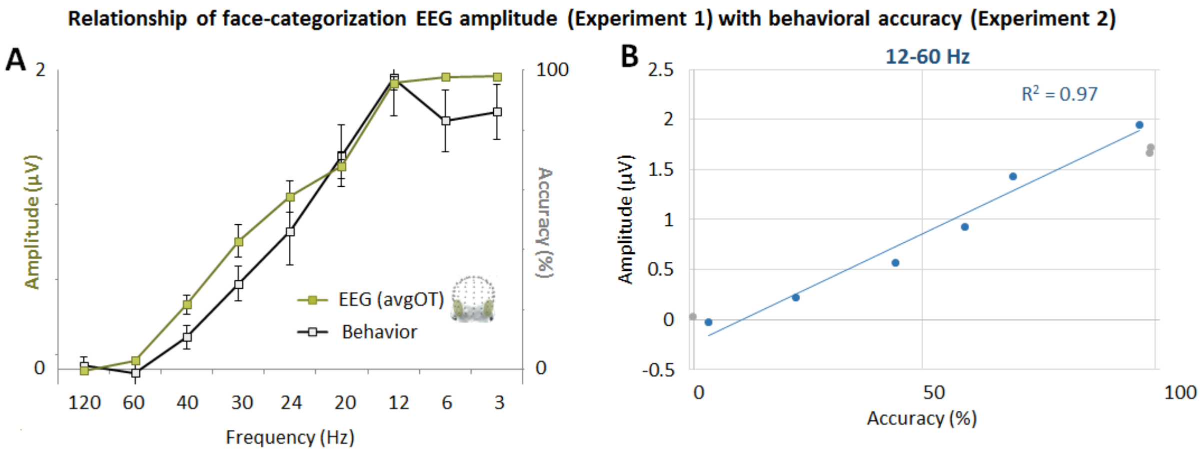


Figure 6. *Behavioral face-categorization accuracy (Experiment 2) compared with neural face-categorization response amplitude (Experiment 1).* ***A)*** *Baseline-subtracted summed-harmonic amplitudes over the OT (avgOT) ROI are plotted (left y-axis scale) along with behavioral accuracy (right y-axis scale).* ***B)*** *A linear relationship between accuracy (Experiment 2) and amplitude (Experiment 1) between 12 and 60 Hz. The data points from the remaining presentation rates (3, 6, and 120 Hz) are shown in light gray, but were not included in this correlation.*

3.2 Behavioral Face Categorization Predicts EEG Face-Categorization Amplitude (Experiment 2)

To further test our hypothesis, we examined whether the neural face-categorization amplitude would be reduced at higher relative to lower rates in the same proportion as the reduction of the behavioral face-categorization accuracy. However, here, we used data quantified in the time domain, within the same experiment (non-periodically presented faces in Experiment 2), focusing on "intermediate" frequency rates, i.e., for which face categorization responses were neither at ceiling nor floor . Specifically, we examined neural responses across two relatively high presentation rates (24 and 30 Hz) with fewer correct detections, and compared them to two relatively low presentation rates (12 and 20 Hz) with more correct detections, for nine participants meeting a minimum behavioral accuracy threshold (see Methods).

In a control comparison, when only behaviorally categorized faces were considered (**Figure 7C**, left panel; see also **Figure 7A; Figure 7B**), there was no significant difference in the combined P1-face and N1-face component amplitudes across 12 and 20 Hz, (M = 3.12 µV; SE = 0.47 µV) vs. 24 and 30 Hz (M = 3.38 µV; SE = 0.72 µV), t_8 = -0.06, p = .48, d = 0.02 faces), Bayes factor of 0.34. However, when behaviorally categorized and non-categorized faces were considered together, the amplitude of the combined amplitude was 30.6% larger at 12 and 20 Hz (M = 2.80 µV; SE = 0.53 µV) than at 24 and 30 Hz (M = 1.95 µV; SE = 0.48 µV), a significant difference, t_8 = 2.29, p = .026, d = 0.56, Bayes factor of 1.41 (**Figure 7C**, right panel; see also **Figure 7A; Figure 7B**). Strikingly, the behavioral accuracy was larger by a similar extent: 32.2% (from 85.8%, SE = 3.27%, to 58.1%, SE = 6.44%) from 12 and 20 Hz to 24 and 30 Hz, for these participants, again a significant difference, t_8 = 6.35, p = .0001, d = 1.80 (**Figure 7D)**. This corresponding decrease of behavioral accuracy and EEG amplitude supports the hypothesis that missed face categorizations contribute proportionally to decreased neural response amplitudes at shorter presentation durations.

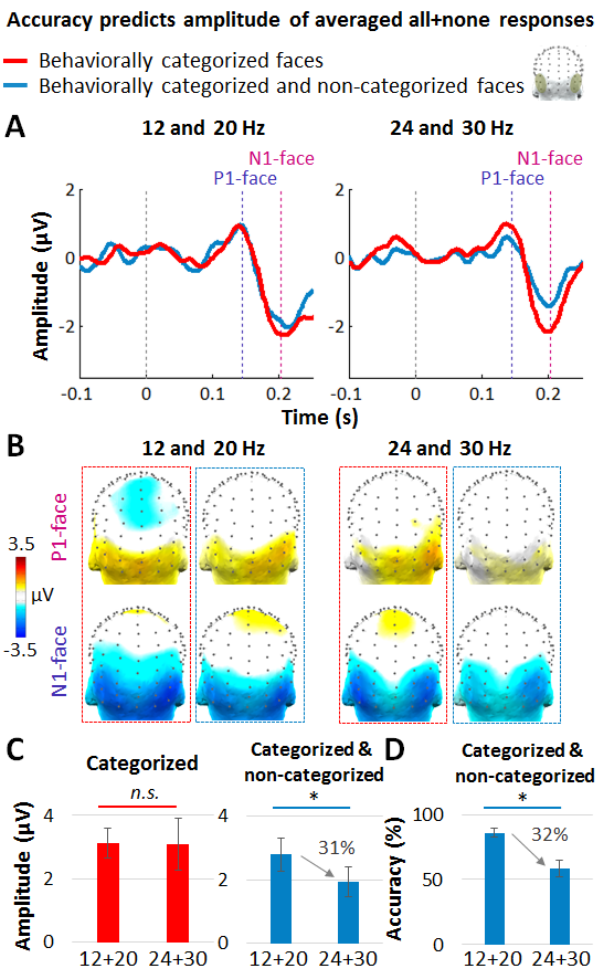


Figure 7. Time-domain responses from Experiment 2 to behaviorally categorized and non-categorized faces together (blue), compared to categorized faces alone (red). There were fewer categorized faces at the higher (24 and 30 Hz), relative to lower (12 and 20 Hz) frequencies; results are shown for participants who had a minimum of 20 artifact-free face-categorizations at both frequency pairs (N = 9). ***A****) Waveforms of the bilateral OT ROI for 12 and 20 Hz (left), and 24 and 30 Hz (right), for both behaviorally categorized faces only (red) and behaviorally categorized and non-categorized faces combined (blue). The indicated times of the P1-face (144 ms) and N1-face (202 ms) components are taken from Experiment 1 (see Figure 6A).* ***B****) Scalp topographies of these two components; outlines indicating condition are colored as in the previous panels.* ***C****) The combined amplitude of the P1-face and N1-face components, at the OT ROI. Responses to behaviorally categorized faces alone did not differ across these conditions (left), but responses to behaviorally categorized and non-categorized faces together were significantly larger at 12 and 20 Hz than at 24 and 30 Hz (right).* ***D****) The percent accuracy dropped from 12 and 20 Hz to 24 and 30 Hz by about 32%; note the correspondence with the 31% decreased EEG amplitude across these conditions in Panel C for to behaviorally categorized and non-categorized faces.*

4. Further Support of All-or-None Neural Responses: Despite Amplitude Differences across Presentation Rates, Categorization Response are Qualitatively Similar

If all-or-none neural responses account for the graded amplitude differences across stimulus presentation rates, we predict that responses are reduced in amplitude at faster presentation rates, but otherwise unchanged in terms of scalp topographies and spatiotemporal dynamics. We investigated these aspects of the face-categorization responses with the robust data from Experiment 1.

4.1 Response Topographies

In support of similar response scalp topographies across presentation rates, the bilateral occipito-temporal (OT) region of interest represented 7 to 9 out of the 10 channels displaying the maximal face-categorization response amplitudes across presentation conditions from 3 to 40 Hz, although only 2 at 60 Hz and 0 at 120 Hz (**Figure 4B**: top row). For further verification of this ROI across presentation rates, we controlled for differences in response amplitude across presentation duration conditions by normalizing the response topographies: this shows the same occipito-temporally dominated responses again from about 40 to 3 Hz (**Figure 4B**: bottom row).

This region was also separated into the right and left hemispheres to assess potential right lateralization effects. In terms of lateralization, the ROT appeared to consistently have a higher amplitude than the LOT, except at 3 Hz. Indeed, the lateralization index (see Methods) revealed a right lateralization (values > 0) across all significant conditions (30 – 3 Hz), except 3 Hz ($M = -3.13$, $SE = 4.33$; see Figure 3A for data at each the right and left OT ROI). The right lateralization index was typically low, ranging from 2.18 to 16.1 across 30 to 6 Hz; it was maximal at 20 Hz (SE = 11.6).[1] Yet, at the group level, there were no significant differences in lateralization across conditions, $F_{1.4,11} = 1.35$, $p = .31$, $\eta_p^2 = 0.38$.

Overall, the response scalp topographies were similar across conditions, being centered over bilateral occipito-temporal channels. This is line with the qualitatively similar (i.e., "all") neural responses being present across different presentation rates, although in different amounts relative to non-categorized faces ("none" neural responses), rather than responses at different stimulus presentation rates having different cortical sources.

4.2 Spatio-Temporal Dynamics

The face-categorization responses in Experiment 1 were examined in the time domain to explore further potential qualitative differences across presentation rates, which would not be predicted by an all-or-none neural response account. The following analyses use data selectively filtering out the stimulus-presentation frequency (see Methods; for unfiltered data, see **Supplemental Figure 4**). Four prominent time-windows of interest appeared across conditions, corresponding to those reported previously as P1-face, N1-face, P2-face, and P3-face (at a 12.5-Hz stimulus-presentation rate; Retter & Rossion, 2016a). These time-windows encompassed nearly all significant deviations over the bilateral OT ROI (**Figure 8A**). Each of these deflections, or "components", was significant across three to six frequency conditions, without any clear trend of component differences as stimulus presentation duration varied. Across significant conditions, these four components peaked over the OT ROI at: 144 ms (SE = 1.95 ms); 202 ms (SE = 5.09 ms); 281 ms (SE = 5.34 ms); and 463 ms (SE = 12.8 ms) post face-stimulus onset.

---

[1] Note that one left-handed participant included in these analyses also presented with a right-lateralized scalp topography (ranked 11/16 in terms of the right lateralization index averaged over 30-3 Hz).

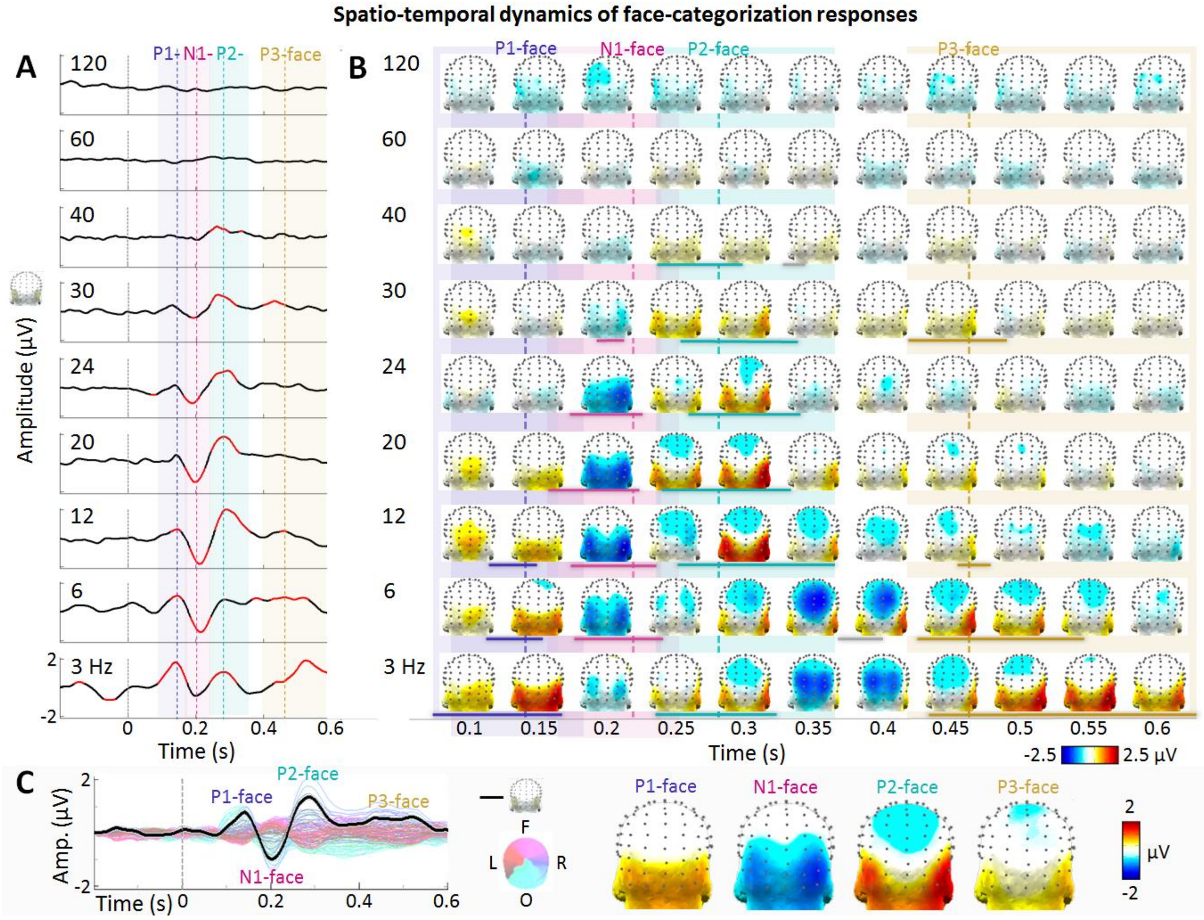


__Figure 8__. Time-domain responses to face presentation from Experiment 1, with each frequency condition filtered to remove the response at its respective stimulus-presentation frequency. Four components were present repeatedly across conditions: P1-face, N1-face, P2-face, and P3-face, and are indicated by a vertical line showing the average peak time, and shading depicting the range, across significant conditions. __A__) Waveforms of the bilateral OT ROI in response to face stimulus onset (0 s). For each condition, significant time periods (p<.001) are plotted in red, while those insignificant are plotted in black. __B__) Scalp topographies, plotted from 0.1-0.6 s post-stimulus onset, every 0.05 s. Significant response periods are underlined in the color of their respective component for each condition; significant periods not corresponding to one of these four components are indicated in gray. __C__) An average of all conditions with significant face-categorization responses in the frequency domain. The bilateral occipito-temporal ROI is plotted in thick black, superimposed above the data from all

128 EEG channels, colored accordingly to the adjacent 2D scalp topography (F: frontal; R: right; O: occipital; L: left). The topographies are plotted to the right at the peaks of the P1-face (142 ms), N1-face (202 ms), P2-face (286 ms), and P3-face (519 ms) for this waveform.

These deflections, and responses more generally across time, occurred mainly over the bilateral OT ROI, with some coverage of the occipito-inferior, occipito-parietal, and medial-occipital regions (**Figure 8B**). In concurrence with the frequency-domain analysis, no significant deflections were found for the conditions at 120 or 60 Hz; however, one significant time-window was found for 40 Hz at the P2-face time, perhaps contributing to the non-significant amplitude in the frequency domain over the OT region for this condition (see again Figure 4B).

In addition to the four main time-windows, a significant time period emerged in the 6 Hz condition between 366 – 398 ms, peaking at 380 ms. This 6-Hz OT positive deflection was accompanied by a large medial parieto-occipital negativity, peaking about 15 ms earlier, which was previously described as the “N2-face” (with a 5.88-Hz sinewave presentation; Jacques, Retter & Rossion, 2016). Although not significant over the OT scalp ROI in other conditions, this response signature is also apparent when stimuli are presented at 3 Hz, and perhaps at 12 Hz and 24 Hz (see again Figure 6B).

Given the overall similarity of responses across presentation rates, a frequency-independent face-categorization response was characterized, based on the average of all conditions with significant face-categorization responses as determined from the frequency domain (3 – 30 Hz) (**Figure 8C**). Over the OT ROI, this response first reached significance at 103 ms post-stimulus onset, and retained significance for 460 ms: it was consistently significant at the times of the P1-face (103 to 159 ms), N1-face (175 to 234 ms), P2- and P3-face (247 to 560 ms). Again, similar spatio-temporal dynamics are in line with qualitatively similar (i.e., “all”) neural responses occurring across different presentation rates, rather than responses being partially degraded at faster presentation rates (e.g., showing reduced or absent amplitudes at selective deflections).

5. Individual Differences in Generic Face Categorization

Finally, we evaluated one more extension of all-or-none neural responses: namely, that across individuals EEG amplitude of face categorization as measured in Experiment 1 correlates with behavioral performance as measured independently in Experiment 2 .

5.1 The Diagnosticity of Presentation Rate (Experiments 1 and 2)

Again, if behavioral face categorization "hits" and "misses" as presentation rate changes can explain changes in amplitude, this should be also relevant at the individual participant level. As a first step, we checked what effect presentation rate had on the range of individual differences, both behaviorally and neurally. To illustrate, the best participant at behavioral face detection (S07) was contrasted to the worst (S01), in terms of percent accuracy and EEG amplitude. (Note that the same participants would have been ranked as the best and worst if accuracy alone was used as a metric.) Large differences between these participants in both behavioral and OT EEG amplitude responses were apparent in the mid-frequency range, from 20 – 30 Hz (50 – 33.3 ms; **Figure 9A**). No differences were apparent across these participants at the lowest three presentation rates (3 – 12 Hz, i.e., 333 – 83.3 ms) in either measure, and differences were greatly reduced by 40 and 60 Hz (25 and 16.7 ms, respectively) until they were abolished at 120 Hz (8.33 ms). The data of all participants, grouped into quartiles, followed these trends, at least between the best 25% and worst 25% of participants (**Figure 9B**).

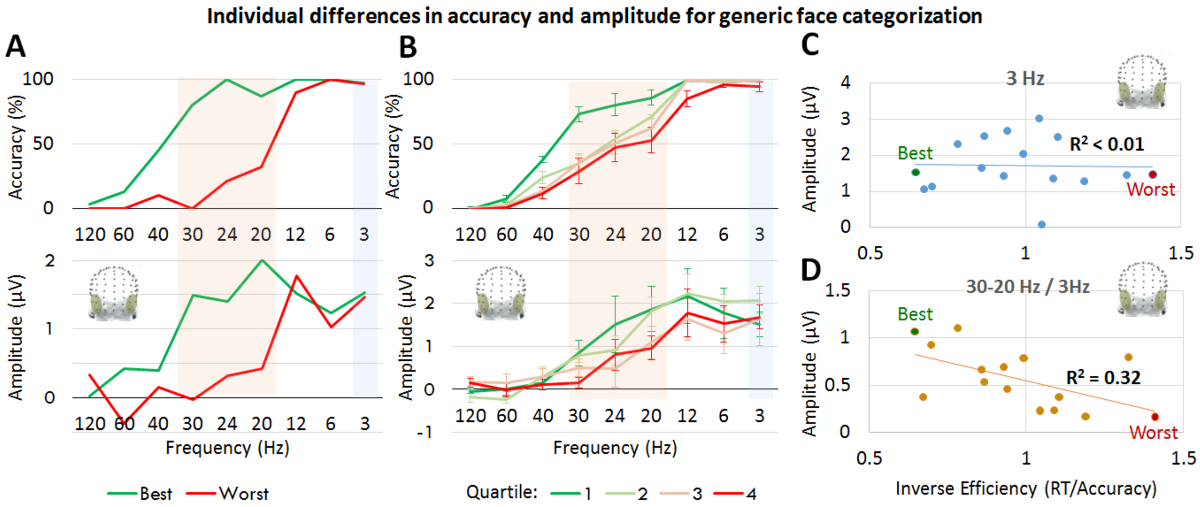


Figure 9*. Individual differences in behavioral performance (Experiment 2) related to EEG amplitude (at the OT ROI; Experiment 1) for face categorization.* ***A)*** *Presentation rates diagnostic of individual differences. A comparison of the participant with the best vs. worst behavioral (inverse efficiency; IE) responses across presentation rates. The highlighted sections indicate data ranges used in panels C and D. Above: Accuracy; Below: Summed baseline-subtracted harmonic amplitude over the bilateral OT ROI.* ***B)*** *As in Part A, except the 16 participants were split into quartiles of four participants each, ranked 1st-4th behaviorally (Group 1), 5th-8th (Group 2), 9th-12th (Group 3), and 13th-16th (Group 4).* ***C)*** *The non-significant correlation of individuals' IE (across all presentation rates) with their OT ROI EEG amplitudes at 3 Hz (highlighted in blue in the previous two panels).* ***D)*** *The significant correlation between individuals' IE with their mid-frequency OT EEG amplitudes (averaged over 20, 24, and 30 Hz, highlighted in orange in the previous two panels), weighted by the amplitude at 3 Hz. In both panels C and D, the data from the participants yielding the best and worst behavioral performance are labeled, respectively.*

The lack of significant EEG responses in the high presentation rates (40 – 120 Hz) would prevent a relationship between individuals' EEG and behavior at this range. At low frequencies (3 – 12 Hz), behavior near ceiling would prevent a relationship between individuals' EEG and behavior. The best separation in EEG response amplitudes was produced at 20 Hz, and behavioral responses seemed well-separated across the mid-frequency range (from 20 – 30 Hz; 50 – 33.3 ms). In the following section, therefore, we will relate individuals' EEG response amplitudes to their behavioral performance using this mid-frequency range.

5.2 The Correlation of Individual Participant's EEG Amplitudes (Experiment 1) and Behavioral Performance (Experiment 2)

Individual participants were ranked for their behavioral face-categorization performance based on their inverse efficiency (IE) score across all presentation rates (range = 0.65 to 1.36; M = 0.97; SE = 0.054; see **Table 1C**), a measure that integrates behavioral performance in terms of both accuracy and response time. To examine whether there was an inherent relationship between individuals' behavioral face-categorization performance (IE scores, computed across presentation rates) and OT EEG face-categorization amplitudes, we correlated these measures at the slowest presentation rate, 3 Hz (333 ms). This correlation was not significant, without any apparent trend, $R^2 = .0006$; $p = 0.93$ (**Figure 9C**), which is likely explained by variable physiological factors influencing the variations of the EEG response amplitude across participants (see Discussion).

When correlating individuals' behavior and amplitudes in the mid-frequency range, from 20 to 30 Hz, as defined in the previous section, the correlation was improved but still not significant, $R^2 = 0.22$; $p = .067$. However, when weighting the individuals' mid-frequency range EEG amplitudes by their baseline amplitudes at 3 Hz, a significant correlation was found, $R^2 = 0.32$, $p = .043$ (**Figure 9D**). This measurement is reflective of the change in amplitude of the EEG response from its amplitude at a presentation rate with ceiling-level behavioral face-categorization performance, and thus likely reflects the percent change in behavioral performance as presentation rate decreases. Yet, note that the weighted correlation was not significantly stronger than the unweighted correlation, $Z = 0.99$, $p = .16$ (one-tailed).

Discussion

With a novel sweep frequency-tagging visual categorization paradigm, we (1) determined the minimal and optimal presentation duration necessary to elicit both behavioral and neural face-categorization responses in human observers, (2) related behavioral and neural responses at the group and individual levels, and (3) determined whether visual categorization is graded or all-or-none. In a first experiment, we presented faces periodically within non-face object RSVP sequences, and investigated implicit response in the EEG frequency domain. In a second experiment, we presented faces non-periodically within non-face object RSVP sequences, and investigated explicit behavioral categorization responses, as well as their associated electrophysiological responses. Note that although faces appeared periodically in Experiment 1 and non-periodically in Experiment 2, comparing neural face categorization responses across experiments is valid because such responses are immune to the periodicity of the stimulus (Quek & Rossion, 2017).

All-or-None Categorization Responses

Our results suggest that stimuli are either categorized as faces and elicit full responses, even at very brief presentation durations (e.g., 17 ms), or they are not categorized, even sometimes at relatively low presentation durations (e.g., 50 ms), and no face-selective responses are elicited. Indeed, across presentation rates from 12 to 60 Hz (83.3 to 17.6 ms), large face-selective EEG deflections were present in response only to behaviorally categorized faces (Figure 5). Additionally, these face-categorization responses obtained at 12 and 20 Hz were not significantly larger than at 24 and 30 Hz when only behaviorally categorized faces were taken into account (Figure 7). In contrast, no significant neural responses to behaviorally non-categorized faces were present in Experiment 2 ($p = .12$); this null result was weakly supported by a Bayes factor of 0.49.

All-or-none categorization responses are not inconsistent with the apparent gradient of response amplitude observed as presentation rate changed in Experiment 1 (Figure 4A), which has been described in many studies varying stimulus visibility (e.g., Kleinschmidt et al., 1998; Bar et al., 2001; Moutoussis & Zeki, 2002; Jemel et al., 2003; Horovitz et al., 2004; Tanskanen et al., 2007; Rousselet et al., 2008; Navajas et al., 2013), including rapid face categorization as measured here but with variable spatial frequency filters (Quek et al., 2018). However, our data indicate that the *proportion* of non-categorized faces drives graded amplitude differences, similar to how behavioral categorization of "hits" and "misses" drives graded behavioral response accuracies. Supporting such an explanation, the behavioral and frequency-tagged neural face-categorization responses were strongly correlated across Experiments 2 and 1, respectively (Figure 6). Furthermore, the decrease in the percent of faces that were categorized behaviorally (a drop of 31% when comparing 24 and 30 Hz to 12 and 20 Hz) strikingly matched the decrease in the amplitude (32%) of the neural face-categorization response as quantified in the same experiment (Figure 7C and D).

It should be noted that the correspondence of neural and behavioral measures as image processing time varies (as a function of presentation duration and/or backward-mask latency) is a replication of a number of previous studies (e.g., Rolls & Tovee, 1994; Kovács, Vogels & Orban, 1995; Vanni et al., 1996; Grill-Spector et al., 2000; Keysers et al., 2001; Bacon-Macé et al., 2005). However, such studies, if interpreted in this light at all, have been taken as evidence for graded perceptual effects, attributing the decrease in response as a linear decrease in the response to each stimulus presented (e.g., see Keysers et al., 2001; Bacon-Mace et al., 2005). Our account and evidence of this relationship as variations in the *proportion* of all-or-none neural responses, is to our knowledge novel and suggests a different underlying process.

Further in line with all-or-none categorization, face-categorization responses remained qualitatively similar in terms of scalp topography (consistent maximal responses over the right occipto-temporal region: Figure 4B) and time course (no trends by presentation duration affecting the P1-face, N1-face, P2-face, or P3-face deflections: Figure 8). Qualitatively similar face-categorization responses in both the frequency and time domain have also been reported with this frequency-tagging categorization paradigm across studies using different stimulus-presentation rates from about 6 to 20 Hz (e.g., Rossion et al., 2015; Retter & Rossion, 2016a; Quek & Rossion, 2017; Retter et al., 2018) and across variations of stimulus visibility (i.e., low-pass filtering, Quek et al., 2018). Moreover, absolute (i.e., against a uniform visual field) responses to faces or other visual objects appear similar both quantitatively and qualitatively when using variable image presentation durations, for example from about 40 to 500 ms (Cichy, Pantazis & Oliva, 2014).

While many arguments have been made for a progressively increasing response amplitude enabling stimulus recognition in the visual cortex (e.g., Kleinschmidt et al., 1998; Bar et al., 2001; Moutoussis & Zeki, 2002; Jemel et al., 2003), all-or-none neural responses have been proposed previously for stimulus detection or identification (e.g., Del Cul, Baillet & Dehaene, 2007; Quiroga et al., 2008; Marti & Dehaene, 2017; see also Navajas et al., 2013; Sekar, Findley & Llinas, 2012; Sekar et al., 2013 for unmasked stimuli). However, the stimuli used in these previous studies are limited in range and well-segmented (e.g., a few letters, numbers or words in Del Cul et al., 2007; Sekar, Findley & Llinas, 2012; Sekar et al., 2013; a few repeated exemplar images in Quiroga et al., 2008; full front segmented faces in Navajas et al., 2013; unknown number of image exemplars in Marti & Dehaene, 2017). Moreover, the timing of these all-or-none neural responses is unclear, with relatively late all-or-none responses often reported (e.g., more than 270 ms in widespread regions in Del Cul et al., 2007; Marti & Dehaene, 2017; more than 300 ms in Quiroga et al., 2008 for neurons in the human medial temporal lobe) as a result of a confound with explicit behavioral reports (see Tsuchiya et al., 2015; Koch et al., 2016). Compared to these studies, our study provides original evidence for all-or-none neural categorization responses, given that: 1) it directly measures categorization (i.e., a differential and generalizable response across widely variable exemplars); 2) through frequency tagging, the neural response is objectively identified and quantified relative to surrounding noise level in the frequency domain; 3) the neural markers of face categorization in this paradigm are known to originate from face-selective neuronal populations in the ventral occipito-temporal cortex (Jonas et al., 2016; Gao et al., 2018); and 4) it shows unambiguous face categorization responses emerging at a 17 ms duration/SOA, and with a latency of 100 ms following stimulus onset.

Finally, the interpretation of all-or-none categorization responses does not imply that at very short presentation durations the perception of a face stimulus is full, i.e., that all the potential visual information, such as age, gender, and identity, has been extracted from the image. Thus, perception may be graded in the sense that a face may be recognized as a face before its identity is recognized. However, we suggest that the generic face categorization process itself is all-or-none: if neural responses differentiating faces from objects and generalizing across face exemplars are elicited at all, they are elicited fully.

Why Are Some Stimuli Not Categorized?

While high-level (i.e., category-selective) visual processes are all-or-none, they may still be reliant on the gradual accumulation of low-level visual information in early visual areas of the brain (e.g., Windey et al., 2014). At an early stage, we theorize that if enough visual information is assembled from the stimulus within about 20 ms, it goes on to develop and trigger a full category-selective response in higher-order brain regions after approximately 100-ms post-stimulus onset (e.g., Crouzet et al., 2010; Crouzet & Thorpe, 2011; Liu et al., 2009; Jacques et al., 2016; Retter & Rossion, 2016a; see Carandini et al., 2005 for a review of this process up to V1). Indeed, in previous studies, the minimum duration to evoke a neural response was between about 12-20 ms when images were backward-masked (e.g., with scrambled images; see Breitmeyer, 1984); this measure was replicated with a wide variety of recording techniques, including single unit in monkeys (Rolls & Tovee, 1994; Kovács, Vogels & Orban, 1995;

Keysers et al., 2001), and MEG (Vanni et al., 1996), fMRI (Grill-Spector et al., 2000), and EEG (Bacon-Macé et al., 2005) in humans. Here, a highly sensitive 20 ms window is consistent with responses occurring rarely when interrupted with a backward-mask after 17 ms, and large increases in the amount of categorization with the next durations (SOAs) of 25 and 30 ms.

However, if early visual processing is interrupted after about 20 ms but before the category-selective response onset (at approximately 100 ms here, and also in Retter & Rossion, 2016a), perceptual categorization may still not occur. Beyond its initiation, visual information persists in the visual system, e.g., for about 60-100 ms, in early (e.g., V1) brain areas (e.g., see the Discussions of Rolls & Tovee, 1994; Keysers et al., 2001). At these stages, the accumulation of visual information may still be disrupted from masking producing competition in early visual areas (see the "neural competition" theory of Keysers and Perrett, 2002; see also the discussion of Potter, 2012). In these experiments, the earlier a mask is presented following a face, the shorter the sensory processing time, and the more likely that the face will not be categorized. In our experiment, missed face categorizations occurred more often at shorter presentation durations, and even up to SOAs of 50 ms (20 Hz; Figure 6A; Figure 3B). At each face presentation, whether or not the image is categorized or not likely depends on many factors, including: properties of the face image (e.g., how central or high-contrast the face is, or the face's head orientation), the overlap of low-level properties with its flanking images (Crouzet & Thorpe, 2011; see again Potter, 2012; but also Maguire & Howe, 2016; Broers et al., 2018), the local context (e.g., whether the preceding face has been categorized), and the neural representation of faces in the stimulated brain.

Note that in this case, a strong effect of masking is not predicted in high-level visual areas, because stimuli temporally flanking the faces consist of non-face objects for at least 1 s before and after (in contrast, see Retter & Rossion, 2016a for a case of high-level face competition when face stimuli are presented with less than about 400 ms SOAs). In previous forward- and backward-masking studies, early competition from masking stimuli was shown to contribute to reduce mean response amplitude when target stimuli were shown for as low as 28 ms; in contrast, with a "gap" between stimuli resulting in SOAs of at least 100 ms, a reduction in amplitude was not produced (Keysers et al., 2005; Retter et al., 2018). Nevertheless, we suggest that if a face is not categorized because of low-level interference before about 100 ms, it does not enter perceptual awareness (as discussed initially by Edelman, 1978, p. 81).

The Minimal Speed for Conscious Face Categorization

A minimum duration of 17 ms (60 Hz presentation rate) was required to elicit occasional face categorization responses in both neural and behavioral measures. This minimal duration is conservative in that it results from highly variable images: the faces varied in size, eccentricity, head orientation, gaze direction, expression, etc. Nevertheless, behavioral face categorization responses were present significantly, but minimally, with 17 ms of image presentation (60 Hz; M = 3.03%, SE = 1.26%, with more frequent responses than false alarms evident at this rate in half (8/16) of the individual participants). In contrast, face categorization was clearly not possible at the shortest presentation duration tested, 8 ms (120 Hz), at which only one participant had only one categorization.

At a neural level, face-categorization responses occurred occasionally at 17 ms over the bilateral occipito-temporal ROI (evident in Experiment 2; data not shown), although these were significant for only one participant in Experiment 1, who had the highest accuracy at this presentation rate. Neural responses were instead not significant at 17 ms for most participants in Experiment 1, likely due to an insufficient number and proportion of detected trials to separate EEG signal from noise: e.g., a participant detecting the average 3% of faces in the second experiment would only have detected about 4 of 126 faces presented in the first experiment at 17 ms. In comparison, the best participant behaviorally detected 12.5% of faces in the second experiment, predicting about 16 faces detected in the first experiment at 17 ms. Note that these 16 responses to faces presumably compete for amplitude with the 110 faces that were not detected, such that recording any significant response is notable, and likely due to the high signal-to-noise ratio afforded from frequency-domain analyses of periodic responses (Norcia et al., 2015).

Perhaps due to such factors, group-level significance was not evident here until 33 ms, when group mean accuracy was 22% (predicting about 28/126 faces detected in the first experiment; Figure 6A; Table 1A). Across participants, the minimum presentation duration required for a significant EEG face-categorization response over the bilateral OT ROI ranged from 17 to 83 ms (Figure 3A). Note that while a linear relation between behavior and neural responses was significant across all presentation rates here, a threshold for recording neural responses predicts a non-linear relationship between behavior and neural responses at the fastest presentation rates (i.e., below about 17 ms here; see Figure 6).

Although very short, the minimum duration of 17 ms for face categorization is consistent with the results of several previous studies, both at behavioral and neural levels. Notably, in a study by Keysers and colleagues (2001; 2005; see also Perrett et al., 2009), the temporal tuning of the responses of single cells and populations in macaque monkeys was investigated in response to monkey heads presented at varying orientations in RSVP sequences. Neural firing rates were shown to decrease at faster presentation rates, but responses to non-periodic target stimuli could still be detected at up to 14 ms (with 300-400 trials), the fastest rate in that experiment. Behavioral accuracy was measured separately with human participants, which also decreased with presentation speed and remained significant up to 14 ms. A later behavioral study using RSVP closely replicated this presentation duration limit for high-level visual perception: Potter et al. (2014) showed recognition of various natural images (e.g., people or flowers) by human participants at 13 ms, the fastest rate in their experiment. Further, additional studies have confirmed the presence of face (and object) detection responses at 16-17 ms with behavior, ECoG, and MEG, in masked or short RSVP sequences (Grill-Spector & Kanwisher, 2005; Fisch et al., 2009; Mohsenzadeh et al., 2018). As mentioned above, these measured durations are likely affected by stimulus factors, such as the types of target stimuli use and the optimization of the stimulus masks (see again Crouzet & Thorpe, 2011; Potter, 2012; Maguire & Howe, 2016; Broers et al., 2018). However, the common limit of approximately 13-17 ms across these studies perhaps reflects something general in visual perception: the minimum amount of time required for a visual stimulus to be coherently processed in early visual areas, before triggering a high-level classification response.

It should be noted that our study was limited in its ability to present stimuli at many durations near potential thresholds, due to the constraints of monitor refresh rates. For example, durations between 17 and 8 ms could not be tested, since these durations consist of two and one screen display frames, respectively, on a 120 Hz monitor. This limitation in available presentation durations has similarly affected many previous studies, and for example limits pinpointing whether a limit is at 13 (Potter et al., 2014), 14 (Keysers et al., 2001), 16 (Fisch et al., 2009), or 17 ms (Grill-Spector & Kanwisher, 2005; Mohsenzadeh et al., 2018). Accordingly, it would also be difficult to assess subtler questions, such as whether faces may be identified faster than other objects around such upper limits (after 17 ms, three display frames at 120 Hz determined the next available duration at 25 ms; but see Fisch et al., 2009; and for a different approach, see Rousselet et al., 2003). Such questions require faster framerates than most current displays afford.

The Optimal Speed of Face-Categorization

Behavioral accuracy neared ceiling at around 80 ms (12 Hz), the presentation rate at which the maximal neural amplitude occurred (Figure 4A). This optimum time aligns well with ranges from previous studies: for example, maximum accuracy and EEG responses were reported at

81 ms for categorizing animal images followed by dynamic visual masks (adjacent durations of 44 and 106 ms; Bacon-Mace et al., 2005). Additionally, a rough plateau of fMRI and behavioral responses between two duration steps of 120 to 500 ms was reported for stimulus naming (reduced responses at the adjacent lower step of 40 ms; see Fig. 3 of Grill-Spector et al., 2000). While we expect variation by the type of visual process targeted, our results suggest that 80 ms is sufficient for neurotypical human adult observers to categorize any natural view of a face in natural viewing conditions (exempting the cases that faces are occluded, artificially degraded, at low contrast, etc.). A practical implication is that studies measuring neural as well as behavioral generic face (vs. object) categorization responses need not present stimuli longer than 80 ms.

Above 80 ms, behavioral accuracy remains at ceiling (above 97% at 6 and 3 Hz; Table 1A). Neural amplitude does not increase at these rates: if anything, it slightly decreases (Figure 4A). In the section "Why Are Some Stimuli Not Categorized?", we discussed how early visual responses may produce masking effects, or competition at a neural level, across neighboring images in RSVP streams. These responses were not specifically represented in the targeted face-categorization responses of the present experiments, which reflect the contrast of faces with non-face objects. However, we also measured general visual responses to stimulus presentation, maximal over the medial occipital cortex, and increasing in amplitude exponentially as stimulus presentation duration increased (Supplementary Figure 1). It is possible these visual responses were more developed as presentation duration increased, being more likely to reach higher-level general visual processing that nevertheless might have interfered more with face-selective processing. Indeed, these stimulus-presentation responses showed some variance in topography across presentation duration, with a particularly distinct response localization at 60 and 120 Hz (Supplemental Figure 3B), although these latter responses may have been contaminated by noise from the power line frequency at 60 Hz. Thus, one possible explanation for lower amplitudes at longer presentation durations could be an interaction between face and non-face object processing at rates below 12 Hz (compare Supplemental Figure 3 with Figure 4). Another, related, explanation, is that harmonics of the 1 Hz face categorization response that overlap with the stimulus presentation frequency at 3 Hz and 6 Hz are excluded in the quantification, underestimating the face categorization response at these frequencies.

Individual Differences in Generic Face Categorization

Finding reliable correlations between neural amplitudes and behavioral measures of visual categorization, even when using frequency-tagging with well quantified measures, remains a challenge (e.g., Retter & Rossion, 2016b; Retter & Rossion, 2017; Xu et al., 2017). This is likely because neural responses are heavily influenced by physiological factors: in particular, for EEG the orientation of dipole generators, influenced by individual differences in cortical folding, is thought to play a large role; additional factors include skull thickness and the conductivity of tissue in between the generators and captors (Luck, 2005; Nunez & Srinivasan, 2006; Woodman, 2010).

Here, we developed a novel method to identify the relationship between individual participants' behavioral and neural measures. We were able to show a significant correlation between participants' behavior (measured in Experiment 2 across all presentation rates in terms of inverse efficiency) and face-categorization EEG amplitude (measured in Experiment 1 by weighting the mid-frequency responses from 20 – 30 Hz (30 – 50 ms) by that of 3 Hz (333 ms), over the OT ROI; Figure 9D).

This correlation relies on two elements: first, we used a range of presentation rates that revealed individual differences in generic face categorization. While a great amount of research is being done in terms of individual differences in face identity recognition (e.g., Wilmer et al., 2014;Yovel et al., 2014; McCaffery et al., 2018), differences in generic face vs. object categorization have not been reported to our knowledge, likely due to the high level of expertise that neurotypical human adults have at performing this function (Crouzet et al., 2010; Herschler & Hochstein, 2005; Scheirer et al., 2014). Nevertheless, here, at presentation rates from 20 to 30 Hz (30 to 50 ms; see Figure 9A and B), we were able to make this function challenging enough to keep all participants below a behavioral ceiling (at 80 ms, i.e., 12 Hz, and later) and yet above the threshold for measuring significant neural responses in most individual participants (Figure 6A; Figure 3).

Second, we used a relative measure of neural activity, weighting individuals' neural amplitudes in this mid-frequency range by their amplitudes at the lowest presentation rate. At this rate, 3 Hz, individual differences had no relationship to behavioral performance (Figure 9C), again likely due to physiological influences on the neural amplitudes. Using a relative neural measure likely assisted in cancelling out such influences, enabling a more diagnostic signature of individuals' EEG response amplitudes relative to behavior (again, Figure 9D).

Note again that this relationship is found when comparing explicit behavioral face categorization (Experiment 2, face detection task) to implicit neural face-categorization responses (Experiment 1, naïve, fixation-cross task). Although most (14/16) participants spontaneously reported perceiving faces when describing the sequences after Experiment 1, none reported observing their periodic presentation. Note also that the responses to stimulus-presentation duration showed very different trends, being significant at all durations but decreasing exponentially as duration increased (Supplemental Figure 3), suggesting that the targeted responses are truly specific to face categorization.

Summary

We have shown that widely variable natural images of faces can be categorized in contrast to non-face objects at presentation rates as short as 17 ms, and optimally at 80 ms. Individual differences were evident within this range, such that individuals' relative neural amplitudes (across 30-50 ms) were significantly correlated with their behavioral performance (across all presentation rates). While overall the amplitude of face-categorization responses appeared to increase gradually as presentation duration increased, this was completely accounted for by the proportion of discrete “hits” and “misses” in categorization, with categorized faces producing full neural face-categorization responses and non-categorized faces producing no neural face-categorization responses. This provides convergent neural and behavioral evidence that human (face) categorization, correlated with perceptual awareness, is not gradual or linear but all-or-none.

Acknowledgments

This research was supported by grants from the Belgian National Foundation for Scientific Research [http://www.fnrs.be/; grant number FC7159 to TR], the European Research Council [https://erc.europa.eu/; grant facessvep number 284025 to BR], the University of Nevada, Reno Integrative Neuroscience Center of Biomedical Research Excellence [https://www.nigms.nih.gov/Research/DRCB/IDeA/pages/COBRE.aspx; grant number P20 GM103650], and the National Eye Institute [https://nei.nih.gov/; grant number EY010834 to MW, and grant number EY023268 to FJ]. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. We are grateful to Andrea Conte for the stimulation codes; to Caroline Michel, for collecting the stimuli; and to Charles Or for generating the color key for plotting electrode location on a scalp map.

References

Adrian E. D., Matthews B. H. C. (1934). The Berger rhythm: Potential changes from the occipital lobes in man. *Brain*, 4 (57), 355–385.

Allison, T., Ginter, H., McCarthy, G., Nobre, A. C., Puce, A. Luby, M. & Spencer, D. D. (1994). Face recognition in human extrastriate cortex. *Journal of Neurophysiology,* 71(2), 821-825.

Bachman, T. (2013). On the all-or-none rule of conscious perception. *Frontiers in Human Neuroscience*, 7, 387.

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.

Bar, M., Tootell, R. B. H., Schacter, D. L., Greve, D. N., Fischl, B., Mendola, J. D. Rosen, B. R. & Dale, A. M. (2001). Cortical mechanisms specific to explicit visual object recognition. *Neuron*, 29(2), 529-535.

Breitmeyer, B. G. (1984). *Visual masking: an integrative approach*. Oxford: Oxford University Press.

Broers, N., Potter, M. C. & Nieuwenstein, M. R. (2018). *Psychon Bull Rev*, 25:1080.

Bruce, V. & Young, A. (1998). In the eye of the beholder: The science of face perception. New York, NY, US: Oxford University Press.

Bushnell I. W. R. (2001) Mother's face recognition in newborn infants: Learning and memory. *Infant and Child Development*, 10, 67–74.

Caharel, S., Leleu, A., Bernard, C., Viggiano, M.P., Lalonde, R., & Rebaï, M. (2013). Early holistic face-like processing of Arcimboldo paintings in the right occipito-temporal cortex: evidence from the N170 ERP component. *Int J Psychophysiol.*, 90, 157-64.

Carandini, M., Demb, J. B., Mante, V., Tolhurst, D. J., Dan, Y, & Olshausen, B. A. et al. (2005). Do we know what the early visual system does? *Journal of Neuroscience*, 25, 10577–10597.

Christensen, M. S., Ramsøy, T. Z., Lund, T. E., Madsen, K. H., & Rowe, J. B. (2006). An fMRI study of the neural correlates of graded visual perception. *Neuroimage*, 31, 1711–1725.

Cichy, R. M., Pantazis, D. & Oliva, A. (2014). Resolving human object recognition in space and time. *Nature Neuroscience*, 17(3), 455-462.

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.

de Gardelle, V., Charles, L., & Kouider, S. (2011). Perceptual awareness and categorical representation of faces: Evidence from masked priming. *Consciousness and Cognition: An International Journal*, 20(4), 1272-1281.

Del Cul, A. Baillet, S. & Dehaene, S. (2007). Brain Dynamics Underlying the Nonlinear Threshold for Access to Consciousness. *PLoS Biology*, 5(10), e260.

Dzhelyova, M. & Rossion, B. (2014). The effect of parametric stimulus size variation on individual face discrimination indexed by fast periodic visual stimuliation. *BMC Neuroscience*, 15:87, 1-12.

Edelman, G. M. (1978). Group selecting and phasic reenterant signalling : A theory a higher brain function. In *The Mindful Brain* (pp. 51-100). Cambridge, MA: MIT Press.

Edelman, G. M. (1987). *Neural Darwinism: The Theory of Neuronal Group Selection*. Cambridge, MA: MIT Press.

Enns, J.T., & Di Lollo V. (2000). What's new in visual masking? *Trends in Cognitive Sciences*, 4, 345-352.

Fisch, L., Privman, E., Ramot, M., Harel, M., Nir, Y., Kipervasser, S. Andelman, F. et al. (2009). Neural “ignition”: enhanced activation linked to perceptual awareness in human ventral stream visual cortex. *Neuron*, 64, 562-574.

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, 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.

Goldstone, R. L., Kersten, A. & Carvalho, P. F. (2018). *Categorization and Concepts*. In Stevens' Handbook of Experimental Psychology and Cognitive Neuroscience, J. T. Wixted (Ed.). doi:10.1002/9781119170174.epcn308

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. & Kanwisher, N. (2005). Visual recognition: as soon as you know it's there, you know what it is. *Psychological Science,* 16(2), 152-160.

Grill-Spector, K., Weiner, K. S., Kay, K. & Gomez, J. (2017). The functional neuroanatomy of human face perception. *Annual Review of Vision Science,* 3, 167-196.

Harris, J. A., Wu, C. T. & Woldorff, M. G. (2011). Sandwich masking eliminates both visual awareness of faces and face-specific brain activity through a feedforward mechanism. *Journal of Vision,* 11(7): 10.

Helmholtz, H. (1867). "*Handbuch der physiologischen Optik*", in Allgemeine Encyklopddie der Physik volume 9, Ed. G Karsten (Leipzig: Voss)

Herrmann, C.S. (2001). Human EEG responses to 1–100 Hz flicker: resonance phenomena in visual cortex and their potential correlation to cognitive phenomena. Experimental Brain Research, 137, 346-353.

Herschler, O. & Hochstein, S. (2005). At first sight: a high-level pop out effect for faces. *Vision Research,* 45(13), 1707-1724.

Horovitz, S.G., Rossion, B., Skudlarski, P., & Gore, J.C. (2004). Parametric design and correlational analyses help integrating fMRI and electrophysiological data during face processing. *Neuroimage,* 22, 1587-1595.

Jacques, C.*, Retter, T.L.*, Rossion, B. (2016). A single glance at a face generates larger and qualitatively different category-selective spatio-temporal signatures than other ecologically-relevant categories in the human brain. *NeuroImage*, 137, 21-33

Jacques, C., Witthoft, N., Weiner, K. S., Foster, B. L., Rangarajan, V., Hermes, D. et al. (2016). Corresponding ECoG and fMRI category-selective signals in human ventral temporal cortex. *Neuropsychologia,* 83, 14-28.

Johnson, M. H., Dziurawiec, S. Ellis, H. & Morton, J. (1991). Newborns' preferential tracking of face-like stimuli and its subsequent decline. *Cognition,* 40(1-2), 1-19.

Jonas, J.*, Jacques, C.*, Liu-Shuang, J., Brissart, H., Colnat-Coulbois, S., Maillard, L., Rossion, B. (2016). A face-selective ventral occipito-temporal map of the human brain with intracerebral potentials. *Proc. Natl. Acad. Sci. USA*, 113, E4088-E4097.

Kanwisher, N., McDermott, J. & Chun, M. M. (1997). The fusiform face area: a module in human extrastriate cortex specialized for face perception. *Journal of Neuroscience*, 17, 4302–4311.

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.

Koch C, Massimini M, Boly M, Tononi G. (2016). Neural correlates of consciousness: progress and problems. Nat Rev Neurosci. 17(5), 307-321.

Kovács, G. Vogels, R. & Orban, G. A. (1995). Cortical correlate of pattern backward masking. *PNAS USA,* 92(12), 5587-5591.

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.

Kleinschmidt, A., Buchel, C., Zeki, S. & Frackowiak, R. S. J. (1998). Human brain activity during spontaneously reversing perception of ambiguous figures. *Proceedings of the Royal Society of London B*, 265:1413.

Leleu, A., Rekow, D., Poncet, F., et al. (2019). Maternal odor shapes rapid face categorization in the infant brain. *Developmental Science*. e12877.

Lee, I. A. & Preacher, K. J. (2013, September). Calculation for the test of the difference between two dependent correlations with one variable in common [Computer software]. Available from http://quantpsy.org.

Liu, H., Agam, Y., Madsen, J. R., & Kreiman, G. (2009). Timing, timing, timing: fast decoding of object information from intracranial field potentials in human visual cortex. *Neuron*, 62(2), 281-90.

Luck, S. J. (2005). *An introduction to the event-related potential technique.* Cambridge, MA: MIT Press.

Maguire, J. F. & Howe, P. D. L. (2016). Failure to detect meaning in RSVP at 27 ms per picture. *Attention, Perception, and Psychophysics*. 78:1405.

Marti, S. & Dehaene, S. (2017). Discrete and continuous mechanisms of temporal selection in rapid visual streams. *Nature Communications,* 8:1955, 1-13.

McCaffery, J. M., Robertson, D. J., Young, A. W. & Burton, A. M. (2018). McCaffery, J. M., Robertson, D. J., Young, A. W., & Burton, A. M. (2018). Individual differences in face identity processing. Cognitive research: principles and implications, 3:21, 1-15.

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.

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.

Moutoussis, K. & Zeki, S. (2002). The relationship between cortical activation and perception investigated with invisible stimuli. *PNAS USA*, 99(14): 9527–9532.

Murphy, G. L. (2002). *The big book of concepts.* Cambridge, MA: MIT Press.

Navajas, J., Ahmadi, M., & Quian Quiroga, R. (2013). Uncovering the mechanisms of conscious face perception: a single-trial study of the n170 responses. *J Neurosci.* 33, 1337-1343.

Navajas, J., Rey, H. G. & Quiroga, R. Q. (2014). Perceptual and contextual awareness: methodological considerations in the search for the neural correlates of consciousness. *Frontiers in Psychology,* 5, 959.

Norcia, A.M., Appelbaum, L.G., Ales, J.M., Cottereau, B., Rossion, B. (2015). The steady-state visual evoked potential in vision research: a review. *Journal of Vision*, 15(6):4, 1-46.

Nunez, P. L. & Srinivasan, R. (2006). *Electric Fields of the Brain.* Oxford: Oxford University Press.

Omer, Y., Sapir, R., Hatuka, Y., & Yovel G. (2019). What Is a Face? Critical Features for Face Detection. *Perception*, 48, 437-446.

Oostenveld, R. & Praamstra, P. (2001). The five percent electrode system for high-resolution EEG and ERP measurements. 112(4), 713-719.

Perrett, D. I. (2009). Seeing the future: Natural image sequences produce “anticipatory” neuronal activity and bias perceptual report. *The Quarterly Journal of Experimental Psychology*, 62(11), 2081-2104.

Potter, M. C. (2012). Recognition and memory for briefly presented scenes. *Frontiers in Psychology*, 3:32.

Potter, M. C. & Levy, E. I. (1969). Recognition memory for a rapid sequence of pictures. *J. Exp. Psychol*, 81, 10–15.

Potter, M.C., Wyble, B., Hagmann, C.E., & McCourt, E.S. (2014). Detecting meaning in RSVP at 13 ms per picture. *Attention, Perception, & Psychophysics*, 76(2), 270-279.

Puce, A., Allison, T., Gore, J. C. & McCarthy, G. (1995). Face-sensitive regions in human extrastriate cortex studied by functional MRI. *Journal of Neurophysiology*, 74(3), 1192-1199.

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.

Quiroga, R. Q., Mukamel, R., Isham, E. A., Malach, R. & Fried, I. (2008). Human single-neuron responses at the threshold of conscious recognition. *PNAS*, 105(9), 3599-3604.

Regan D. (1989). *Human brain electrophysiology: Evoked potentials and evoked magnetic fields in science and medicine*. Amsterdam, the Netherlands: Elsevier.

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. & 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. & Rossion, B. (2016b). Visual adaptation provides objective electrophysiological evidence of facial identity discrimination. *Cortex*, 80, 35-50.

Rolls, E. T. & Tovée, M. J. (1994). Processing Speed in the Cerebral Cortex and the

Neurophysiology of Visual Masking. *Proceedings of the Royal Society B,* 257, 9-15.

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.

Rouder, J. N., Speckman, P. L., Sun, D., Morey, R. D. & Iverson, G. (2009). Bayesian t-tests for accepting and rejecting the null hypothesis. *Psychonomic Bulletin and Review*, 16, 225-237.

Rousselet, G. A., Macé M. J.-M, Fabre-Thorpe, M. (2003). Is it an animal? Is it a human face? Fast processing in upright and inverted natural scenes. *Journal of Vision*, 3(6):5.

Rousselet, G.A., Pernet, C.R., Bennett, P.J., & Sekuler, A.B. (2008). Parametric study of EEG sensitivity to phase noise during face processing. *BMC Neurosci.* 9:98.

Scheirer, W. J., Anthony, S. E., Nakayama, K. & Cox, D. D. (2014). Perceptual annotation: Measuring human vision to improve computer vision. *IEE Transactions on Pattern Analysis and Machine Intelligence,* 36(8), 1679-1686.

Sekar, K., Findley, W. M. & Linas, R. R. (2012). Evidence for an all-or-none perceptual response: single-trial analyses of magnetoencephalography signals indicate an abrupt transition between visual perception and its absence. *Neuroscience,* 206, 167-182.

Sekar, K., Findley, W. M., Poeppel, D., & Llinás, R. R. (2013). Cortical response tracking the conscious experience of threshold duration visual stimuli indicates visual perception is all or none. *Proceedings of the National Academy of Sciences of the United States of America*, 110(14), 5642-7.

Sergent, J., Ohta, S. & MacDonald, B. (1992). Functional neuroanatomy of face and object processing. A positron emission tomography study. *Brain*, 115(1), 15-36.

Shafto J. P., Pitts M. A. (2015). Neural signatures of conscious face perception in an inattentional blindness paradigm. *Journal of Neuroscience,* 35, 10940–10948.

Smith, E. E. & Medin, D. L. (1981). *Categories and concepts.* Cambridge, MA: Harvard University Press.

Sugden, N. A., Mohamed-Ali, . I. & Moulson, M. C. (2014). I spy with my little eye: Typical, daily exposure to faces documented from a first-person infant perspective. *Developmental Psychobiology,* 56(2), 249-261.

Tanskanen, T., Näsänen, R., Ojanpää, H., & Hari, R. (2007). Face recognition and cortical responses: effect of stimulus duration. *Neuroimage*, 35, 1636-1644

Thorpe, S. J., Fize, D. & Marlot, C. (1996). Speed of processing in the human visual system. *Nature,* 381(6582), 520-522.

Tsuchiya, N., Wilke, M., Frässle, S. & Lamme, V. A. (2015). No-report paradigms: extracting the true neural correlates of consciousness. *Trends in Cognitive Science,* 19, 757–770

Townsend, J. T. & Ashby, F. G. (1983). *The Stochastic Modeling of Elementary Psychological Processes*. Cambridge: Cambridge University Press.

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.

Weiner, K. S. & Grill-Spector, K. (2010). Sparsely-distributed organization of face and limb activations in human ventral temporal cortex. *Neuroimage,* 52(4): 1559–1573.

Wilmer, J. B., Germine, L. T. & Nakayama, K. (2014). Face recognition: a model specific ability. *Frontiers in Human Neuroscience,* 8:769, 1-5.

Windey, B. & Cleeremans, A. (2015). Consciousness as a graded and an all-or-none phenomenon: A conceptual analysis. *Consciousness and cognition,* 35, 185-191.

Windey, B., Vermeiren, A., Atas, A. & Cleeremans, A. (2014). The graded and dichotomous nature of visual awareness. *Phil. Trans. R. Soc. B,* 369:20130282.

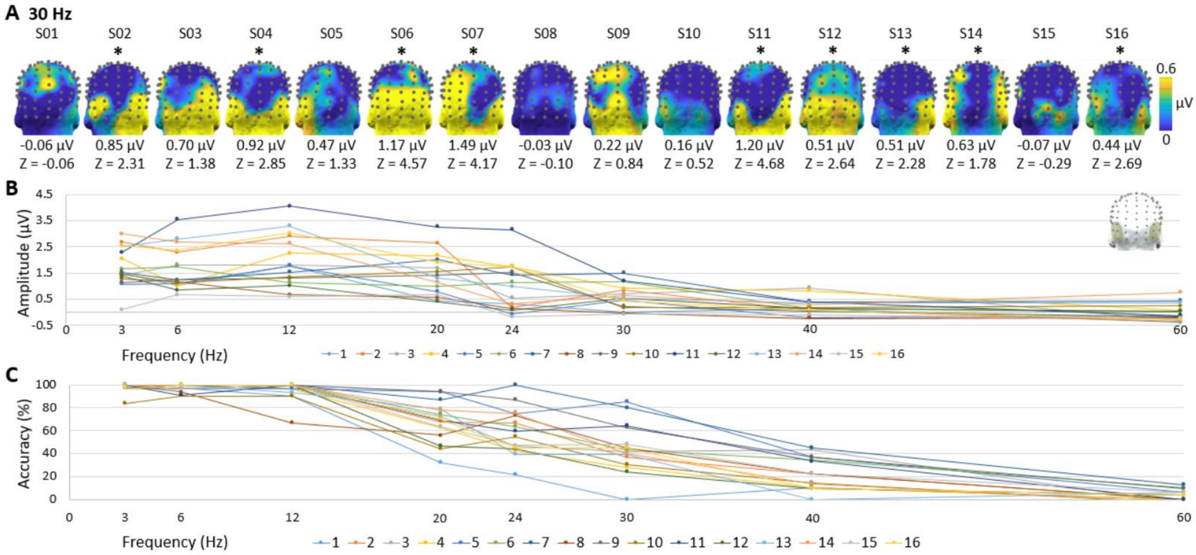
Woodman, G. F. (2010). A brief introduction to the use of event-related potentials in studies of perception and attention. *Attention, Perception, and Psychophysics.* 72(8), 2031-2046.

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.

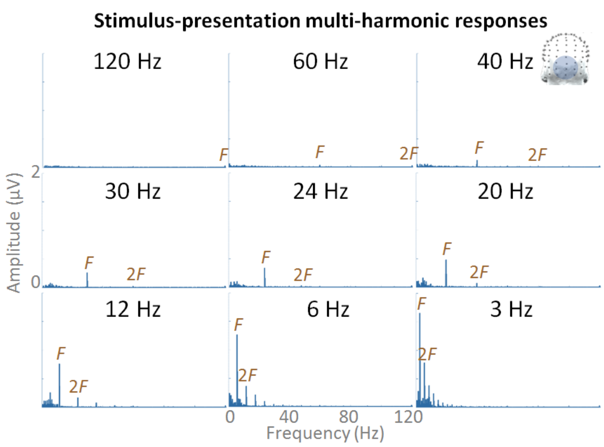
Yovel, G., Wilmer, J. B. & Duchaine, B. (2014). What can individual differences reveal about face processing? Frontiers in Human Neuroscience, 8(562), 1-9.

Zhen, Z., Yang, Z., Huang, L., Kong, X.-z., Wang, X., Dang, X., et al. (2015). Quantifying interindividual variability and asymmetry of face-selective regions: a probabilistic functional atlas. *Neuroimage,* 113, 13-25.

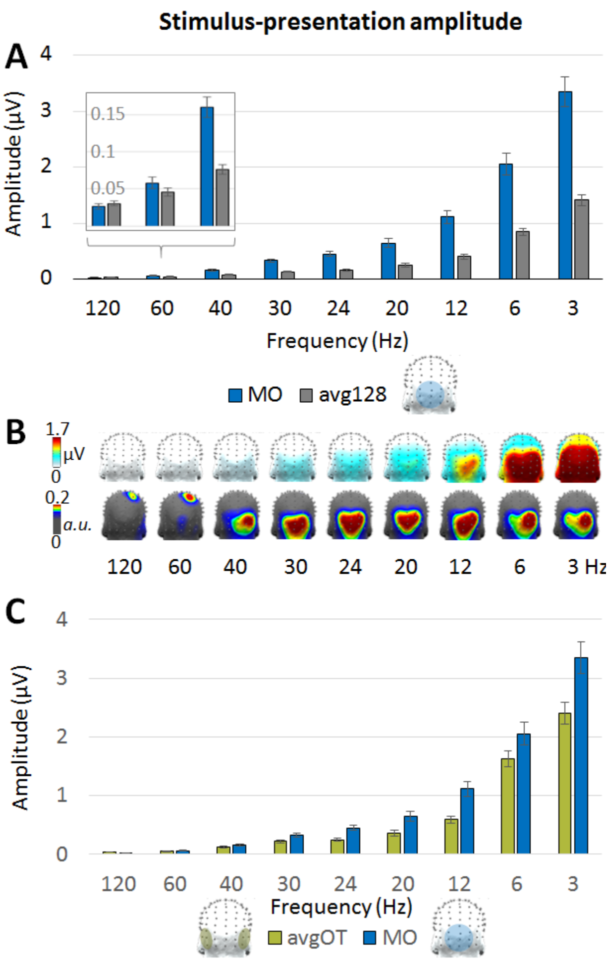
Supplementary Material



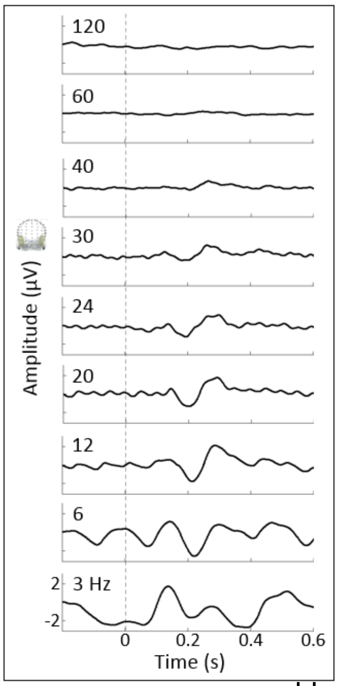
Supplemental Figure 1. *Individual participant data (S01-S16).* ***A)*** *Scalp topographies for the maximal condition producing a significant group-level response (33 ms; 30 Hz). Below each scalp topography is that participant's baseline-*subtracted*, summed-harmonic, bilateral occipito-temporal (OT) ROI amplitude and Z-Score. Nine participants with a significant individual-level response at this frequency are indicated with an asterisk.* ***B)*** *Amplitudes over the bilateral OT ROI across stimulus presentation duration rates up to 60 Hz.* ***C)*** *Accuracies across stimulus presentation duration rates.*



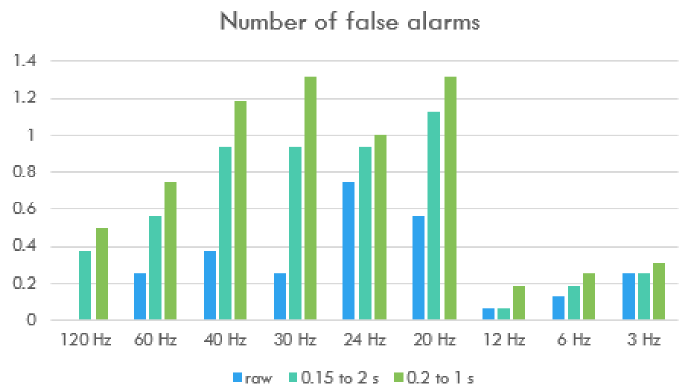
Supplemental Figure 2*. Stimulus-presentation responses for each condition, characterized in the frequency domain in the form of baseline-subtracted amplitude spectra. Stimulus-presentation responses were tagged at the rate of stimulus presentation, F (= 120 Hz, 60 Hz, etc.) in every condition, with higher harmonics occurring at 2F, 3F, etc. (the first two harmonics are labeled above the spectra). The data are shown over* *the region of the scalp (MO ROI) and frequency range (up to1 20 Hz) that were selected to capture stimulus-presentation responses. For ease of comparison across condition, all graphs are plotted with common axes.*



Supplemental Figure 3. *Stimulus-presentation responses as a function of stimulus-presentation rate (Experiment 1), to be compared with Figure 3.* ***A)*** *The middle occipital ROI and average of all 128 channels across presentation rates.* ***B)*** *Scalp topographies. Top row: grand-averaged amplitudes. Bottom row: the same responses normalized.* ***C)*** *The OT and MO ROIs for stimulus-presentation responses.* ***Key)*** *MO = middle occipital; avg128 = average of all 128 EEG channels; avgOT = average of the 10 occipito-temporal channels.*



Supplemental Figure 4. *Stimulus-presentation unfiltered time-domain data by condition, with labeling conventions as in Figure 5A. Note that in processing the data were low-pass filtered at 30 Hz, preventing apparent responses at 120-40 Hz, although the amplitude of stimulus-presentation responses was very low at these rates anyway (see Supplemental Figure 2 and Supplemental Table 2B).*



Supplemental Figure 5. *The mean number of false alarms, summed across all sequence repetitions per frequency (in total about 30 face presentations per frequency). Raw = no restrictive time window for correct responses; note that a window of 0.15 to 2 s was used in the analyses, but a more restrictive window of 0.2 to 1 s would have only slightly increased the rate of false positives across frequencies.*

Frequency (Hz)	120	60	40	30	24	20	12	6	3
A. Face									
OT	0.11	-0.14	1.12	**3.50**	**5.09**	**6.77**	**10.37**	**8.87**	**6.92**
Avg128	0.25	0.04	0.25	0.55	1.46	**2.37**	**3.02**	**3.23**	**2.38**
B. Stimulus									
MO	**14.73**	**24.33**	**49.19**	**72.08**	**50.22**	**98.41**	**71.87**	**328.60**	**63.68**
Avg128	**18.16**	**18.68**	**22.51**	**53.25**	**30.01**	**74.36**	**47.62**	**144.46**	**35.99**

Supplemental Table 1. *Group-level Z-scores (Experiment 1), calculated at the occipito-temporal (OT) ROI for face-categorization responses (**A**), the medial-occipital (MO) ROI for stimulus-presentation responses (**B**), and the average of all 128 EEG channels for both face- and stimulus-presentation responses. Significant responses are shown in bold (Z > 2.32, p<.01).*

Frequency (Hz)	120	60	40	30	24	20	12	6	3
A. Face (µV)									
OT	0.03 (0.06)	-0.03 (0.08)	0.21 (0.08)	0.57 (0.12)	0.93 (0.23)	1.43 (0.21)	1.95 (0.25)	1.66 (0.21)	1.72 (0.19)
MO	0.008 (0.055)	0.07 (0.07)	0.22 (0.09)	0.26 (0.08)	0.47 (0.09)	0.78 (0.13)	1.06 (0.18)	1.05 (0.12)	1.10 (0.18)
B. Stimulus (µV)									
OT	0.04 (0.005)	0.05 (0.006)	0.13 (0.01)	0.22 (0.02)	0.24 (0.03)	0.35 (0.05)	0.59 (0.06)	1.62 (0.13)	2.40 (0.18)
MO	0.03 (0.004)	0.06 (0.008)	0.16 (0.01)	0.33 (0.03)	0.45 (0.04)	0.64 (0.08)	1.11 (0.12)	2.05 (0.19)	3.35 (0.27)

Supplemental Table 2. *Mean response amplitude (µV), calculated at the occipito-temporal (OT) and medial-occipital (MO) ROIs for both face-categorization (**A**) and stimulus-presentation (**B**). One standard error, across participants, is included in parentheses.*