



**HAL**  
open science

## Rapid neural categorization of facelike objects predicts the perceptual awareness of a face (face pareidolia)

Diane Rekow, Jean Yves Baudouin, Renaud Brochard, Bruno Rossion, Arnaud Leleu

### ► To cite this version:

Diane Rekow, Jean Yves Baudouin, Renaud Brochard, Bruno Rossion, Arnaud Leleu. Rapid neural categorization of facelike objects predicts the perceptual awareness of a face (face pareidolia). *Cognition*, 2022, 222, pp.105016. 10.1016/j.cognition.2022.105016 . hal-03542375

**HAL Id: hal-03542375**

**<https://hal.inrae.fr/hal-03542375>**

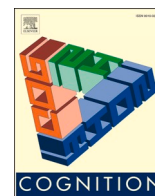
Submitted on 25 Jan 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 International License



## Rapid neural categorization of facelike objects predicts the perceptual awareness of a face (*face pareidolia*)

Diane Rekow<sup>a,\*</sup>, Jean-Yves Baudouin<sup>b</sup>, Renaud Brochard<sup>a</sup>, Bruno Rossion<sup>c,d</sup>, Arnaud Leleu<sup>a,\*</sup>

<sup>a</sup> Laboratoire Éthologie Développementale et Psychologie Cognitive, Centre des Sciences du Goût et de l'Alimentation, Université Bourgogne Franche-Comté, CNRS, Inrae, AgroSup Dijon, F-21000 Dijon, France

<sup>b</sup> Laboratoire Développement, Individu, Processus, Handicap, Éducation (DIPHE), Département Psychologie du Développement, de l'Éducation et des Vulnérabilités (PsyDÉV), Institut de psychologie, Université de Lyon (Lumière Lyon 2), 69676 Bron, cedex, France

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

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

### ARTICLE INFO

#### Keywords:

Visual categorization  
Perceptual awareness  
Face pareidolia  
EEG frequency-tagging

### ABSTRACT

The human brain rapidly and automatically categorizes faces vs. other visual objects. However, whether face-selective neural activity predicts the subjective experience of a face – *perceptual awareness* – is debated. To clarify this issue, here we use *face pareidolia*, i.e., the illusory perception of a face, as a proxy to relate the neural categorization of a variety of facelike objects to conscious face perception. In Experiment 1, scalp electroencephalogram (EEG) is recorded while pictures of human faces or facelike objects – in different stimulation sequences – are interleaved every second (i.e., at 1 Hz) in a rapid 6-Hz train of natural images of nonface objects. Participants do not perform any explicit face categorization task during stimulation, and report whether they perceived illusory faces post-stimulation. A robust categorization response to facelike objects is identified at 1 Hz and harmonics in the EEG frequency spectrum with a facelike occipito-temporal topography. Across all individuals, the facelike categorization response is of about 20% of the response to human faces, but more strongly right-lateralized. Critically, its amplitude is much larger in participants who report having perceived illusory faces. In Experiment 2, facelike or matched nonface objects from the same categories appear at 1 Hz in sequences of nonface objects presented at variable stimulation rates (60 Hz to 12 Hz) and participants explicitly report after each sequence whether they perceived illusory faces. The facelike categorization response already emerges at the shortest stimulus duration (i.e., 17 ms at 60 Hz) and predicts the behavioral report of conscious perception. Strikingly, neural facelike-selectivity emerges exclusively when participants report illusory faces. Collectively, these experiments characterize a neural signature of face pareidolia in the context of rapid categorization, supporting the view that face-selective brain activity reliably predicts the subjective experience of a face from a single glance at a variety of stimuli.

### 1. Introduction

Humans are very good and fast at categorizing visual stimuli as faces (e.g., Crouzet, Kirchner, & Thorpe, 2010; Hershler, Golan, Bentin, & Hochstein, 2010; Scheirer, Anthony, Nakayama, & Cox, 2014). At the neural level, this critical face categorization function is subtended by a large network of cortical areas in the ventral occipito-temporal cortex (VOTC) (e.g., Gao, Gentile, & Rossion, 2018; Jonas et al., 2016; Sergent, Ohta, & Macdonald, 1992; Zhen et al., 2015; Grill-Spector, Weiner, Kay, & Gomez, 2017 for review) and leads to specific signatures in scalp electroencephalography (EEG) (e.g., Bentin, Allison, Puce, Perez, &

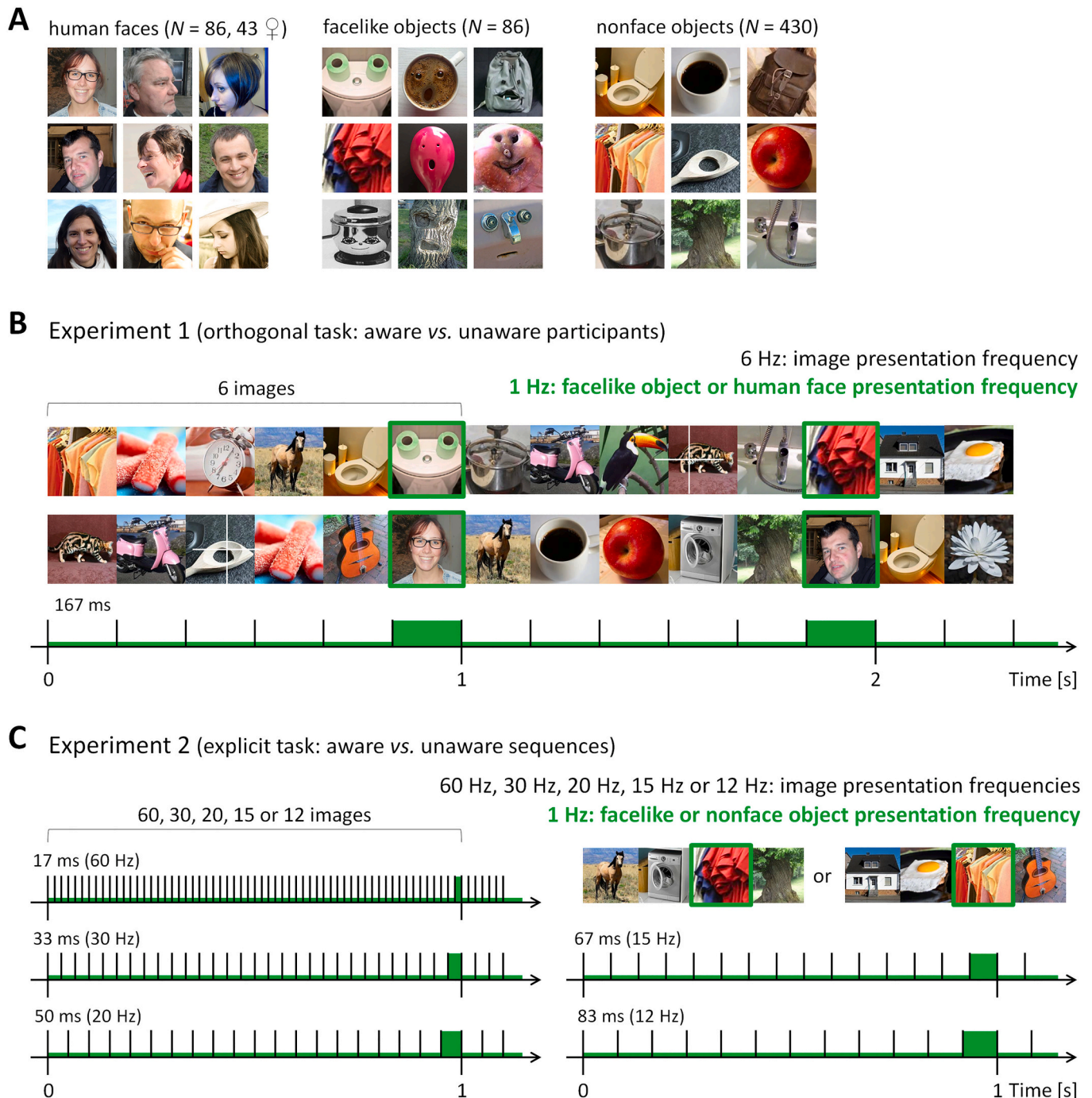
McCarthy, 1996; Jeffreys, 1996; Rossion, Torfs, Jacques, & Liu-Shuang, 2015). However, whether neural face categorization predicts the subjective experience of a face – *perceptual awareness*, is debated (Aru et al., 2012; Harris, Wu, & Woldorff, 2011; Moutoussis & Zeki, 2002; Navajas, Ahmadi, & Quiroga, 2013; Perry, 2016; Philiastides & Sajda, 2006; Retter, Jiang, Webster, & Rossion, 2020; Tanskanen, Näsänen, Ojanpää, & Hari, 2007; Tong, Nakayama, Vaughan, & Kanwisher, 1998). This is essentially due to the challenge of measuring this function in the brain, i.e., measuring a neural response that incorporates high selectivity to faces (vs. many nonface categories) and generalizability across a wide range of variable face stimuli (e.g., Rossion, 2014 for a discussion).

\* Corresponding author at: Centre des Sciences du Goût et de l'Alimentation, 9E Boulevard Jeanne d'Arc, 21000 Dijon, France.

E-mail addresses: [diane.rekow@u-bourgogne.fr](mailto:diane.rekow@u-bourgogne.fr) (D. Rekow), [arnaud.leleu@u-bourgogne.fr](mailto:arnaud.leleu@u-bourgogne.fr) (A. Leleu).

Recently, a valid measure of rapid and automatic face categorization has been developed in scalp EEG (e.g., Jacques, Retter, & Rossion, 2016; Retter & Rossion, 2016; Rossion et al., 2015), neuroimaging (Gao et al., 2018), and intracranial recordings (Hagen et al., 2020; Jonas et al., 2016) using a frequency-tagging approach (Norcia, Appelbaum, Ales,

Cottareau, & Rossion, 2015 for review). With this approach, a rapid stream of forward- and backward-masked natural images of many living and non-living objects is presented at a base rate (e.g., 6 Hz) and variable exemplars of human faces are interspersed at a lower rate (e.g., 1 Hz) while participants do not explicitly categorize faces (Fig. 1A & 1B). This



**Fig. 1.** Measuring face pareidolia with EEG frequency-tagging. **A.** Examples of variable natural images of human faces, facelike objects and nonface objects used as stimuli in both experiments. **B.** Example of  $\approx 2$  s of visual stimulation at a 6-Hz presentation frequency (i.e., 6 images per second without inter-stimulus interval, stimulus duration  $\approx 167$  ms) used in Experiment 1. Facelike objects or human faces are inserted at 1 Hz (i.e., every 6th stimulus) in dedicated sequences. This procedure tags two brain responses in the EEG frequency spectrum: a general response (6 Hz and harmonics, i.e. integer multiples) to all visual cues rapidly changing at 6 Hz, and a categorization response (1 Hz and harmonics) reflecting the visual categorization of facelike objects or human faces (i.e., discrimination from nonface objects and generalization across exemplars). Participants perform a cross-detection task and are asked post-stimulation if they noticed illusory faces to compare perceptually aware and unaware participants. **C.** Examples of  $\approx 1$  s of stimulation at 5 presentation frequencies (60 Hz, 30 Hz, 20 Hz, 15 Hz, 12 Hz; stimulus durations: 17 ms, 33 ms, 50 ms, 67 ms, 83 ms) used in Experiment 2. Facelike or nonface objects are interspersed at 1 Hz in dedicated sequences. This time, participants explicitly attend to the stimuli and report after each sequence if they perceived illusory faces to compare sequences associated with awareness or not.

paradigm thus dissociates two brain responses automatically elicited at predefined frequencies: a general visual response (base rate) and a face categorization response (face presentation rate). The general response reflects the neural activity elicited by the fast train of stimuli, while the face categorization response captures face-selective activity generalized across face stimuli. This latter response is large, sensitive, reliable, and not accounted for by the amplitude spectrum of the images (Gao et al., 2018; Rossion et al., 2015). Albeit sporadic (i.e., recorded for only a subset of faces), the face categorization response is already fully elicited (i.e., of full magnitude) at very high speed of stimulation (i.e., at 60 Hz; stimulus duration: 17 ms) and saturates (i.e., becomes systematic) at 12 Hz (stimulus duration: 83 ms; Retter et al., 2020). Importantly for our purpose, these variable occurrences of the response as a function of stimulus duration are associated with behavioral reports of face perception, revealing that rapid and all-or-none face categorization in the brain predicts perceptual awareness (Retter et al., 2020).

Despite the large variability of the face stimuli used in frequency-tagging studies, they all show clear human facial features, are all recognized as faces by human observers, and very likely to be recognized as faces with high accuracy by an artificial system (e.g., Scheirer et al., 2014; see Grill-Spector, Weiner, Gomez, Stigliani, & Natu, 2018 for a discussion). However, an in-depth understanding of conscious face perception from rapid neural categorization must consider that face percepts also emerge from a variety of inputs in the natural visual environment despite the absence of human facial features, namely *face pareidolia*, one of the most remarkable examples of ubiquitous illusory perception in the human species (see examples in Fig. 1A). Pareidolia is generically defined as “the tendency to perceive a specific, often meaningful image in a random or ambiguous visual pattern” (Merriam-Webster, nd), but more than 70% of pictures represent a face when searching “pareidolia” on the web (estimation made in January 2021 with the first 100 different pictures in Google Images). Face pareidolia has long been used by painters (e.g. Giuseppe Arcimboldo, 1527–1593) or photographers (e.g., Robert & Robert, 1996), and even recently observed in nonhuman primates (Taubert et al., 2018; Taubert, Wardle, Flessert, Leopold, & Ungerleider, 2017; Taubert, Wardle, & Ungerleider, 2020). Perceiving a face in a nonface stimulus leads to several facelike processing advantages, such as improved individuation abilities (Vuong et al., 2017), or higher saliency in a visual search task (Keys, Taubert, & Wardle, 2021).

In the human brain, prior studies have documented how facelike stimuli elicit activity within face-selective regions in the VOTC (Dolan et al., 1997; Hadjikhani, Kveraga, Naik, & Ahlfors, 2009; Kanwisher, Tong, & Nakayama, 1998; McKeef & Tong, 2007; Rossion, Dricot, Goebel, & Busigny, 2011; Wardle, Taubert, Teichmann, & Baker, 2020), or a facelike EEG response over right occipito-temporal scalp sites (Caharel et al., 2013; Churches, Baron-Cohen, & Ring, 2009; Sagiv & Bentin, 2001). Facelike neural activity is generally identified when stimuli are consciously reported as faces by human observers (Andrews & Schluppeck, 2004; Bentin, Sagiv, Mecklinger, Friederici, & von Cramon, 2002; George, Jemel, Fiori, Chaby, & Renault, 2005; McKeef & Tong, 2007; Shafto & Pitts, 2015), contrary to stimuli judged as facelike by a computational face-detection system (Moulson, Balas, Nelson, & Sinha, 2011). Nevertheless, as for human face categorization, previous studies are limited in their ability to unequivocally relate the perceptual awareness of an illusory face to neural facelike categorization for several reasons. *First*, only one type of facelike stimuli and one type of nonface stimuli are generally used with a few different exemplars for each, preventing a valid measure of facelike categorization, i.e., a highly selective response to facelike stimuli (vs. a large set of nonface stimuli) that generalizes across many variable facelike exemplars. In addition, this cannot rule out the contribution of responses elicited by the stimuli irrespective of their category (e.g., responses to low-level cues) to the observed relationship between neural activity and perceptual awareness. *Second* and relatedly, stimuli are often displayed in isolation and controlled for their irrelevant visual properties (e.g., luminance,

viewpoint) by elimination or homogenization (i.e., absent or constant across stimuli). Yet, removing the background and homogenizing stimuli artificially increases facelikeness by delineating a global face shape, and confounds facelikeness with low physical differences across stimuli (see Davidenko, Remus, & Grill-Spector, 2012 for a discussion). *Third* and finally, previous studies have rarely used meaningful objects as facelike stimuli, and when they have, facelike objects were not contrasted to other objects from the same categories to carefully control for shared neural responses between facelike and nonface stimuli.

A recent study has circumvented these issues by using a variety of naturalistic facelike objects together with matched nonface objects and human faces (Wardle et al., 2020). Using both functional imaging and magnetoencephalography, this study used a multivariate classification approach to dissociate facelike and object-like neural activities within the absolute brain response to facelike objects, and revealed an early (i.e., within the first 250 ms after stimulus-onset) facelike response originating from face-selective regions in the VOTC. However, despite its obvious qualities, this study has not clarified whether the brain activity selectively elicited by facelike objects predicts the conscious report of illusory faces in individual participants (i.e., face pareidolia).

To fill this gap in knowledge, here we employ EEG frequency-tagging in two experiments to provide a neural categorization response reflecting the conscious perception of a face in a large set of naturalistic facelike stimuli contrasted to nonface stimuli depicting similar objects (Fig. 1A). The nonface stimuli are presented at the base rate and the facelike stimuli at a lower rate, such that a facelike categorization response emerges only if stimuli depicting similar objects (i.e., facelike and nonface stimuli) elicit dissimilar neural activity, whereas stimuli depicting dissimilar objects (i.e., facelike stimuli) elicit similar neural activity according to their facelikeness (Fig. 1B & 1C). In other words, with this approach, we isolate a direct differential neural response to facelike objects (i.e., without post-hoc subtraction) devoid of shared neural activities between facelike and nonface stimuli. In *Experiment 1*, we present 27 participants with 40-s-long sequences at a 6-Hz base rate (i.e., 6 images per second, stimulus duration  $\approx$  167 ms) and facelike objects or human faces are interleaved every 6 stimuli (i.e., at 1 Hz; Fig. 1B) to tag the categorization of illusory and human faces at 1 Hz and harmonics (i.e., integer multiples) and estimate the facelikeness of the former. Importantly, participants perform an orthogonal cross-detection task ensuring implicit exposure to facelike stimuli. They are then queried post-stimulation whether they noticed facelike objects, and classified as perceptually aware or unaware participants, to determine whether face pareidolia is associated with the automatic neural categorization of facelike stimuli. In *Experiment 2*, another 22 participants are presented with 16-s-long sequences at 5 different base rates (60 Hz, 30 Hz, 20 Hz, 15 Hz, 12 Hz), such that stimulus duration varies from 17 to 83 ms as in Retter et al. (2020). Facelike or nonface objects are always interspersed at 1 Hz in dedicated sequences (i.e., facelike objects in half of the sequences). Contrary to Experiment 1, participants are informed of the presence of facelike objects before testing and must report after each sequence if they perceived illusory faces to determine, for each participant, whether the facelike categorization response unfolds only in sequences associated with perceptual awareness. Overall, through these two experiments, we demonstrate that a facelike categorization response to a wide range of facelike stimuli already emerges at a short 17-ms stimulus duration in the human brain, and predicts conscious illusory face perception in individual participants.

## 2. Materials and methods

### 2.1. Experiment 1

#### 2.1.1. Participants

Twenty-seven participants (12 females, 6 left-handed (3 females), mean age:  $22.5 \pm 2.9$  (SD) years, range: 19–31 years) took part in the experiment and were compensated for their participation. All reported



normal/corrected-to-normal visual acuity, and none reported a history of neurological/psychiatric disorder. They provided written informed consent prior to the experiment. Testing was conducted in accordance with the Declaration of Helsinki and approved by the local ethics committee of the “Université Bourgogne-Franche-Comté” (CERUBFC-2021-08-17-022).

Since participants performed an orthogonal task, we asked them three questions post-stimulation to determine whether they perceived facelike objects. First, we asked if they noticed something particular during the experiment. If they did not mention illusory faces, we then asked whether they noticed something about the stimuli. Note that all participants reported here the presence of human faces but none detected their periodicity. Again, if participants did not mention illusory faces, we finally questioned them about the presentation of facelike objects. Based on this interview, participants were split in two groups, one group that mentioned illusory faces in at least one question (i.e., perceptually *aware* participants,  $N = 13$ , 5 females, 2 left-handed (1 female), mean age:  $23.2 \pm 3.5$  years, range: 19–31 years), and another group that did not (i.e., *unaware* participants,  $N = 14$ , 7 females, 4 left-handed (2 females), mean age:  $21.9 \pm 2.3$  years, range: 19–27 years). The two groups did not significantly differ in age ( $T_{25} = 1.21$ ,  $p = .24$ ), sex ( $X^2_1 = 0.55$ ,  $p = .36$ ), and handedness ( $X^2_1 = 0.41$ ,  $p = .68$ ).

### 2.1.2. Stimuli

Stimuli were color natural images of 86 human faces (43 females), 86 facelike objects and 430 nonface objects cropped to a square and sized to  $300 \times 300$  pixels. All stimuli were embedded in their original scenes and differed in size, viewpoint, lighting and background so that their physical characteristics were widely variable (examples in Fig. 1A, full set available upon request from the authors). In addition, human faces varied largely in age, sex, race and expression. Face and nonface images were adapted from previous studies (e.g., Jacques et al., 2016; Retter & Rossion, 2016; Rossion et al., 2015) or collected from the Internet. Nonface objects were various biological and manufactured objects with several exemplars (i.e., between 3 and 20) in each category (listed in *Supplementary Materials and Methods*). Facelike images were selected among a large set of 224 pictures collected from the Internet when searching for ‘face pareidolia’. Selection was made according to the images judged as the most facelike in a pretest (*Supplementary Materials and Methods* & Fig. S1). Critically, facelike images depicted various object categories (between 1 and 5 exemplars in each category) matching some of those used for nonface objects (listed in *Supplementary Materials and Methods*). Hence, facelike objects differed from nonface objects only in their overall facelike appearance (Fig. 1A).

Face and facelike stimuli were both divided in two sets of 43 pictures. For human faces, one set contained 22 females and the other one 21 females. For facelike images, at least one exemplar of each object category was allocated to each set. These two sets ensured that all face and facelike stimuli were presented to every participant. During the experiment, stimuli were displayed at the center of a 24-in. LED screen (60 Hz refresh rate, resolution:  $1920 \times 1080$  pixels) on a mid-level grey background (i.e., 128/255 in greyscale). From a viewing distance of 57 cm, they subtended approximately  $8.3^\circ$  of visual angle.

### 2.1.3. Procedure

The procedure was adapted from previous face categorization experiments using EEG frequency-tagging (e.g., Jacques et al., 2016; Retter & Rossion, 2016; Rossion et al., 2015). Images were presented at a fast base rate of 6 Hz (i.e., 6 images per second,  $\approx 167$  ms per image) without inter-stimulus interval (forward- and backward-masking; Fig. 1B). In each stimulation sequence, nonface objects were used as base stimuli. In different sequences, human faces or facelike objects were periodically inserted every 6th stimulus (i.e., at  $6/6 = 1$  Hz; 1 s between two human faces or facelike objects).

After electrode-cap placement, participants were seated in a light- and sound-isolated cabin in front of the stimulation screen. Their head

was maintained on a chinrest at a distance of 57 cm from the screen. Stimulation sequences started with a 2-s fade-in of increasing contrast modulation depth (0 to 100%), followed by the full-contrast stimulation lasting 40 s and then followed by a 2-s fade-out of decreasing contrast modulation depth (100 to 0%). Both fade-in and fade-out were used to reduce eye-blinks and movements elicited by the sudden onset or offset of flickering stimuli. Sequences were flanked by variable pre- and post-stimulation intervals of 0.5–1.5 s of uniform grey background. For both face and facelike stimuli, each set of 43 images was used in half of the stimulation sequences while the 430 nonface objects were used in all sequences. Each experimental condition (i.e., category at 1 Hz) was repeated 6 times (i.e., 3 times for each stimulus set), resulting in 12 sequences throughout the experiment. They were divided in 3 blocks of 4 sequences, each block presenting two sequences per condition (i.e., one per stimulus set). Blocks and sequences within blocks were randomly presented across participants. In each sequence, stimuli were randomly selected from their respective sets.

### 2.1.4. Orthogonal behavioral task

An orthogonal behavioral task was designed to ensure that participants were exposed to the stimulation. During each sequence, they were asked to detect 8 brief (200 ms) random appearances of a  $300 \times 300$  pixels large white cross on the images by pressing the spacebar of a keyboard with both index fingers as quickly as possible. A minimum interval of 2 s was introduced between two cross-onsets. Both accuracy and RTs for correct detections (ranging between 100 and 1000 ms) were submitted to a repeated-measures ANOVA with *Category* (human faces vs. facelike objects) as a within-subject factor and *Group* (aware vs. unaware) as a between-subject factor.

### 2.1.5. EEG acquisition and preprocessing

Scalp electroencephalogram (EEG) was continuously acquired from a 64-channel BioSemi Active-Two amplifier system (BioSemi, The Netherlands) with Ag/AgCl electrodes located according to the 10–10 classification system. During recording, the Common Mode Sense (CMS) active electrode was used as reference and the Driven Right Leg (DRL) passive electrode was used as ground. Electrode offset was held below  $\pm 15$   $\mu$ V for each electrode and EEG was sampled at 1024 Hz.

All EEG analyses were carried out in Letswave 6 (<https://www.lets-wave.org/>) running on Matlab 2017 (MathWorks, USA). For each participant, continuous datasets were first bandpass filtered at 0.1–100 Hz using a Butterworth filter (4th order) and then downsampled to 256 Hz. Datasets were segmented into 45-s epochs for each stimulation sequence (12 per participant, 2 conditions  $\times$  6 repetitions), including 1 s before the fade-in and 1 s after the fade-out. An Independent Component Analysis (ICA) with a square mixing matrix was computed (Makeig, Bell, Jung, & Sejnowski, 1996) to isolate and remove components corresponding to eye-blinks (i.e., one component recorded over Fp channels per participant) and to additional artifacts recorded over frontal and temporal channels (mean number across participants: 2.3, range: 0–4, no significant difference between groups of participants,  $T_{25} = 0.91$ ,  $p = .37$ ). Remaining noisy or artifact-ridden channels were replaced using linear interpolation from the 4 neighboring channels (mean number across participants: 0.9, range: 0–5, no significant difference between groups of participants,  $T_{25} = 0.16$ ,  $p = .88$ ). EEG epochs were then re-referenced to the average of the 64 channels.

### 2.1.6. EEG frequency-domain analysis

In line with previous face categorization studies (e.g., Jacques et al., 2016; Retter & Rossion, 2016; Rossion et al., 2015), our paradigm was designed to tag two different brain responses at two predefined frequencies within a single stimulation sequence, and to quantify them in the EEG amplitude spectrum using frequency-domain analysis: (1) a general response at 6 Hz and harmonics (i.e., integer multiples) elicited by the stream of images (i.e., both nonface and face/facelike images) and capturing the visual response to all cues (e.g., local contrast) rapidly

changing 6 times per second; (2) a categorization response at 1 Hz and harmonics reflecting the *differential* response to face or facelike stimuli. Thanks to the rapid and periodic mode of stimulation, this response indexes single-glance visual categorization of human faces and facelike objects implying discrimination from nonface objects and generalization across category exemplars despite widely variable images. It is not accounted for by the amplitude spectrum of the stimuli (Gao et al., 2018; Rossion et al., 2015) and is immune to temporal predictability elicited by periodicity (Quek & Rossion, 2017). Note that the 1-Hz rate of face or facelike presentation allows enough time between image-onsets (i.e., 1 s) for the full face categorization response to develop ( $\approx 450$  ms in duration, Retter & Rossion, 2016).

For each participant, the 6 epochs recorded for each condition were averaged to reduce EEG activity non-phase-locked to the stimuli, thus resulting in a single 45-s epoch per condition. Epochs were then precisely cropped from the onset of the full-contrast stimulation to 40 s so as to contain an exact integer number of 1-Hz cycles (i.e., 40 cycles, 10,240 time bins). A fast Fourier transform (FFT) was applied to every epoch and amplitude spectra were extracted for all channels with a high frequency resolution of  $1/40 = 0.025$  Hz. Thanks to this high resolution, 40 frequency bins were extracted every 1-Hz step, allowing unambiguous identification of the tagged brain responses and estimation of noise amplitude from surrounding frequency bins. Given our objective to identify a selective response to facelike objects that reflects their categorization as faces, we considered the EEG data recorded for sequences containing human faces as a reference to determine the range of harmonics (i.e., tagged frequencies and their integer multiples) and regions-of-interest (ROIs) for further analysis.

The range of harmonics (Table S1) for the brain responses to the 6-Hz stimulation and the 1-Hz face presentation was defined from the FFT amplitude spectra averaged across channels and participants (Fig. 2A). Z-scores were calculated as the difference between the amplitude at the target frequency bin and the mean amplitude of the surrounding noise ( $\approx \pm 0.3$  Hz: estimated from 20 frequency bins, 10 on each side, excluding the immediately adjacent and the 2 most extreme (minimum and maximum) bins) divided by the standard deviation of the noise. Harmonics were included until Z-scores were no longer significant (i.e.,  $Z > 1.64$ ,  $p < .05$ , one-tailed, signal > noise). For the general response, all harmonics were significant (i.e., 8 harmonics, from 6 Hz to 48 Hz, harmonics were not considered beyond the 50 Hz response elicited by AC power). For the face categorization response, harmonics reached significance up to 26 Hz (i.e., 26th harmonic). The overall responses were then condensed by summing amplitudes across significant harmonics (excluding harmonics corresponding to the general response (i.e., 6 Hz, 12 Hz, 18 Hz, 24 Hz) for the categorization response) for each category, channel and participant. Henceforth, mentions of the general and categorization responses to either human or illusory faces will refer to these amplitudes summed across harmonics.

To quantify the magnitude of each brain response in a value expressed in microvolts ( $\mu\text{V}$ ), we isolated the response from noise level by subtracting out the mean amplitude of the surrounding frequency bins, leading to notional corrected amplitudes of zero in the absence of response. Corrected amplitudes were used to define the ROIs, and the ROIs were used to conduct group-level statistical analyses and significance estimation of individual brain responses. The strength of each brain response within ROIs (Fig. 2B & 2C) was also estimated with signal-to-noise ratios (SNR) computed by dividing the uncorrected response amplitude by the mean surrounding noise.

In line with previous studies (Jacques et al., 2016; Rossion et al., 2015), both categorization and general responses in the human face condition present a right hemisphere advantage, but the categorization response is laterally distributed over occipito-temporal regions (Fig. 2B) while the general response is located over the middle occipital cortex (Fig. 2C). Accordingly, we defined two symmetrical occipito-temporal ROIs for the face categorization response and one middle occipital ROI for the general visual response by considering the channels with the

maximal group-level corrected amplitudes. For the face categorization response, the largest amplitude was observed over channel P10 (4.27  $\mu\text{V}$ ), followed by PO8 (3.81  $\mu\text{V}$ ), PO7 (3.37  $\mu\text{V}$ ), P8 (3.27  $\mu\text{V}$ ), P9 (2.96  $\mu\text{V}$ ) and P7 (2.93  $\mu\text{V}$ ). We thus defined homologous right and left occipito-temporal ROIs (respectively rOT and lOT), each comprising 3 contiguous channels (r/lOT: P10/9, PO8/7, P8/7). For the general visual response, the strongest amplitudes were recorded over channels O2 (3.50  $\mu\text{V}$ ), PO8 (3.32  $\mu\text{V}$ ), Oz (3.10  $\mu\text{V}$ ) and O1 (3.04  $\mu\text{V}$ ). The single middle occipital ROI thus encompassed these 4 neighboring channels.

Statistical analyses were conducted separately for the categorization and general responses. In addition, for each response, two separate analyses were also consecutively performed. The first one evaluated the response elicited by facelike objects compared to human faces with the whole participant sample for an initial characterization of the neural signature of illusory face perception. We ran a repeated-measures ANOVA on individual corrected amplitudes with *Category* (human faces vs. facelike objects) as a within-subject factor. The within-subject factor *Hemisphere* (rOT vs. lOT) was also included for the categorization response. The second analysis aimed at determining whether the neural patterns identified in the first analysis depend on the perceptual awareness of illusory faces. For this purpose, repeated-measures ANOVAs were conducted for each stimulus category with *Group* (aware vs. unaware) as a between-subject factor. In all analyses, post-hoc comparisons were conducted for significant effects using *t*-tests and the false discovery rate (FDR) procedure was applied to adjust *p*-values for multiple comparisons (Benjamini & Hochberg, 1995). Since corrected amplitudes should not differ from zero in the absence of response, significance of the grand-averaged brain responses was estimated by identifying whether the 95% confidence interval (CI, calculated across participants) around the mean response amplitude did not include zero.

Finally, two other analyses were carried on to estimate the significance of the brain responses in every individual participant and to determine whether the topographies of the categorization responses to facelike objects and human faces are reliable and comparable. For the first analysis, the significance of individual responses within ROIs was estimated using Z-scores (see above). For the second analysis, the 6 epochs (i.e., time series) recorded for sequences presenting human faces or facelike objects were split according to stimulus sets, resulting in  $2 \times 3$  epochs for each condition. Epochs were then averaged and following processing steps were similar to those previously described in order to isolate both general and categorization responses to either human faces or facelike objects expressed in summed corrected amplitudes separately for each stimulus set. After grand-averaging individual responses, we computed Pearson's correlations between the categorization responses obtained for each set using the 64 channels as observations. We thus estimated the topographical reliability of the categorization response across stimulus sets for both categories. Correlations were also calculated between both categorization responses to determine whether their scalp distributions are close. As a control index expected to reveal a lower topographical similarity, the correlation between the categorization response and the general response recorded for sequences containing facelike objects was finally computed.

## 2.2. Experiment 2

### 2.2.1. Participants

We tested 22 participants (15 females, 1 left-handed (female), mean age:  $21.4 \pm 4$  (*SD*) years, range: 18–33 years) who did not participate in Experiment 1. They reported normal/corrected-to-normal visual acuity and no history of neurological/psychiatric disorder. They provided written informed consent prior to the experiment and were compensated. Testing was conducted in accordance with the Declaration of Helsinki and approved by the local ethics committee of the “Université Bourgogne-Franche-Comté” (CERUBFC-2021-08-17-022).

### 2.2.2. Stimuli

Stimuli were the 80 facelike objects judged as the most facelike in the pretest conducted before Experiment 1 (*Supplementary Materials and Methods & Fig. S1*) and the 430 nonface objects used in Experiment 1 (examples in *Fig. 1A*). An additional set of 15 facelike images (judged as the most facelike after the 80 first ones) was also used before testing to illustrate which kind of stimuli participants must detect (see section 2.2.4). The 80 facelike stimuli used for testing were divided in five sets of 16 pictures. Nonface objects were divided in one set of 350 stimuli always used as base stimuli, and five sets of 16 stimuli matched to the facelike stimuli in terms of the depicted object categories for sequences containing only nonface objects (see section 2.2.3). Stimulus resolution and size, screen parameters, and viewing distance were identical to Experiment 1.

### 2.2.3. Procedure

The procedure was adapted from a recent EEG frequency-tagging study investigating face categorization at various stimulus durations (*Retter et al., 2020*). Images were presented without inter-stimulus interval (forward and backward-masking) at five stimulation frequencies depending on the sequence: 60 Hz, 30 Hz, 20 Hz, 15 Hz and 12 Hz (i.e., stimulus-onset asynchrony = stimulus duration: 17 ms, 33 ms, 50 ms, 67 ms, 83 ms; *Fig. 1C*). These frequencies were chosen according to the screen refresh rate (i.e., 60 Hz), such that stimulus durations were 1, 2, 3, 4 or 5 frames. In every sequence, nonface objects were used as base stimuli. Facelike objects were interspersed at 1 Hz in half of the sequences. Thus, for instance, at 60 Hz, facelike objects appeared every 60 images, while at 12 Hz they appeared every 12 images (*Fig. 1C*). Facelike stimuli were replaced by the matched nonface objects (see section 2.2.2) in the other half of the sequences. This led to 10 conditions: 2 categories (facelike objects or nonface objects)  $\times$  5 stimulus durations (17 ms, 33 ms, 50 ms, 67 ms, 83 ms; corresponding to 5 stimulation rates: 60 Hz, 30 Hz, 20 Hz, 15 Hz, 12 Hz).

Stimulation sequences started with a 0.5-s pre-stimulation interval, followed by a 1.833-s fade-in of increasing contrast modulation depth (0 to 100%). Then, the full-contrast stimulation lasted 15.167 s before a 1-s fade-out of decreasing contrast modulation depth (100 to 0%) and a 0.5-s post-stimulation interval. For both facelike and nonface objects interleaved at 1 Hz, each set of 16 images was used in half of the stimulation sequences. The 350 nonface objects used as base stimuli were presented in all sequences. Each experimental condition was repeated 10 times, resulting in 100 sequences throughout the experiment. They were divided in 10 blocks of 10 sequences, each block presenting one sequence per condition. Blocks and sequences within blocks were randomly presented across participants. Stimuli were randomly selected from their respective sets.

### 2.2.4. Explicit behavioral task

Contrary to Experiment 1, perceptual awareness of illusory faces was expected to vary for each participant, as a function of stimulus duration (i.e., stimulation frequency). Hence, participants were explicitly instructed to attend to the stimuli and to detect facelike objects. After electrode-cap setup, participants were told that rapid sequences of natural images depicting objects will be presented at variable rates and that they will have to report orally after each sequence whether it contained some objects resembling faces. Fifteen facelike images were presented one by one for illustration (not used thereafter). Participants were informed that some sequences will contain several facelike objects and some will not. They were also informed that because images will be presented at rapid rates, false alarms could be frequent. Accordingly, they were asked to report illusory face perception if and only if they noticed several facelike exemplars throughout the sequence. The number of facelike reports after a sequence (out of 10) was submitted to a repeated-measures ANOVA with *Category* (facelike objects vs. nonface objects) and *Stimulus duration* (17 ms vs. 33 ms vs. 50 ms vs. 67 ms vs. 83 ms) as within-subject factors. Note that this dependent variable

corresponds to the number of hits and false alarms for sequences presenting facelike objects and nonface objects, respectively. Mauchly's test for sphericity violation was performed and Greenhouse-Geisser correction (epsilon:  $\epsilon$ ) for degrees of freedom was applied whenever sphericity was violated. Post-hoc comparisons were conducted using *t*-tests and the FDR procedure was applied to adjust *p*-values (*Benjamini & Hochberg, 1995*).

### 2.2.5. EEG acquisition and preprocessing

EEG acquisition and preprocessing steps were identical to Experiment 1, except for data segmentation into 20-s epochs for each stimulation sequence (100 per participant, 10 conditions  $\times$  10 repetitions), including 1 s before the fade-in and 1 s after the fade-out. Following ICA, the mean number of removed components across participants was 2.7 (range: 1–5). The mean number of interpolated channels was 0.5 (range: 0–3).

### 2.2.6. EEG frequency-domain analysis

As in Experiment 1, the facelike categorization response was tagged at 1 Hz and harmonics. In contrast, contrary to Experiment 1, the general visual response was tagged at different frequencies depending on the sequence (i.e., 60 Hz, 30 Hz, 20 Hz, 15 Hz, 12 Hz and their respective harmonics). For each participant, the 10 epochs recorded for each condition were averaged in the time domain, leading to a single 20-s segment per condition. Epochs were cropped from the onset of the full-contrast stimulation to 16 s (4096 time bins). An FFT was applied and amplitude spectra were extracted with a frequency resolution of  $1/16 = 0.0625$  Hz, leading to 16 frequency bins every 1-Hz step.

For the categorization response, harmonics were included up to the 26th harmonic (i.e., 26 Hz) according to Experiment 1. For the general visual response elicited at variable frequencies up to 60 Hz, we considered harmonics until this frequency for each condition. In other words, only the first harmonic (i.e., 60 Hz) was included for the general response to a 60 Hz-stimulation stream, two harmonics (i.e., 30 Hz and 60 Hz) for the response to a 30 Hz-stimulation, three harmonics (i.e., 20 Hz, 40 Hz, 60 Hz) for a 20 Hz-stimulation, four harmonics (i.e., 15 Hz, 30 Hz, 45 Hz, 60 Hz) for a 15 Hz-stimulation and five harmonics (i.e., 12 Hz, 24 Hz, 36 Hz, 48 Hz, 60 Hz) for a 12 Hz-stimulation. The overall responses were summed across harmonics, excluding those corresponding to the stimulation frequencies and their harmonics (i.e., 12 Hz, 15 Hz, 20 Hz, 24 Hz) for the categorization response. The general and categorization responses will refer to these summed amplitudes thereafter.

The magnitude of each brain response was quantified in mean corrected amplitudes within the ROIs defined in Experiment 1 (i.e., right and left occipito-temporal [r/IOT: P10/9, PO8/7, P8/7] for the categorization response and middle occipital [PO8, O2, Oz, O1] for the general response). To estimate the baseline noise level in a similar frequency range as in Experiment 1 ( $\approx \pm 0.3$  Hz), we considered the mean amplitude of 6 surrounding frequency bins (3 on each side), excluding the adjacent bins and the most extreme (minimum and maximum) bins. Statistical analyses were carried on individual corrected amplitudes separately for the categorization and general responses using repeated-measures ANOVAs with *Category* (facelike objects vs. nonface objects) and *Stimulus duration* (17 ms vs. 33 ms vs. 50 ms vs. 67 ms vs. 83 ms) as within-subject factors. The factor *Hemisphere* (rOT vs. lOT) was also included for the categorization response. Mauchly's test for sphericity violation was performed and Greenhouse-Geisser correction (epsilon:  $\epsilon$ ) for degrees of freedom was applied whenever sphericity was violated. Post-hoc comparisons were conducted using *t*-tests with FDR-adjusted *p*-values (*Benjamini & Hochberg, 1995*). As in Experiment 1, significance of the grand-averaged brain responses was estimated by determining whether the 95% *CI* around the mean amplitude did not include zero. Grand-averaged corrected amplitudes were normalized by their scalp-wide power (*McCarthy & Wood, 1985*) to illustrate the topography of each brain response regardless of its magnitude.

To evaluate the relationship between the behavioral report of illusory faces and the amplitude of the facelike categorization response, we computed Pearson's correlations between individual data for both measures and for each stimulus duration. We first divided each measure by its value at the longest duration (i.e., 83 ms) to correct for individual differences in ceiling-level responses (Retter et al., 2020). In other words, this measurement is free from the between-subject variability observed when performance is at ceiling. We then calculated correlations for the four remaining durations. We computed the same correlation for the average of the two shortest durations (i.e., 17 and 33 ms), which lead to mid-level amplitude of the categorization response and mid-level number of facelike reports, and did the same for the two (i.e., 50 and 67 ms) following durations, which lead to larger neural responses and near-ceiling behavioral performance.

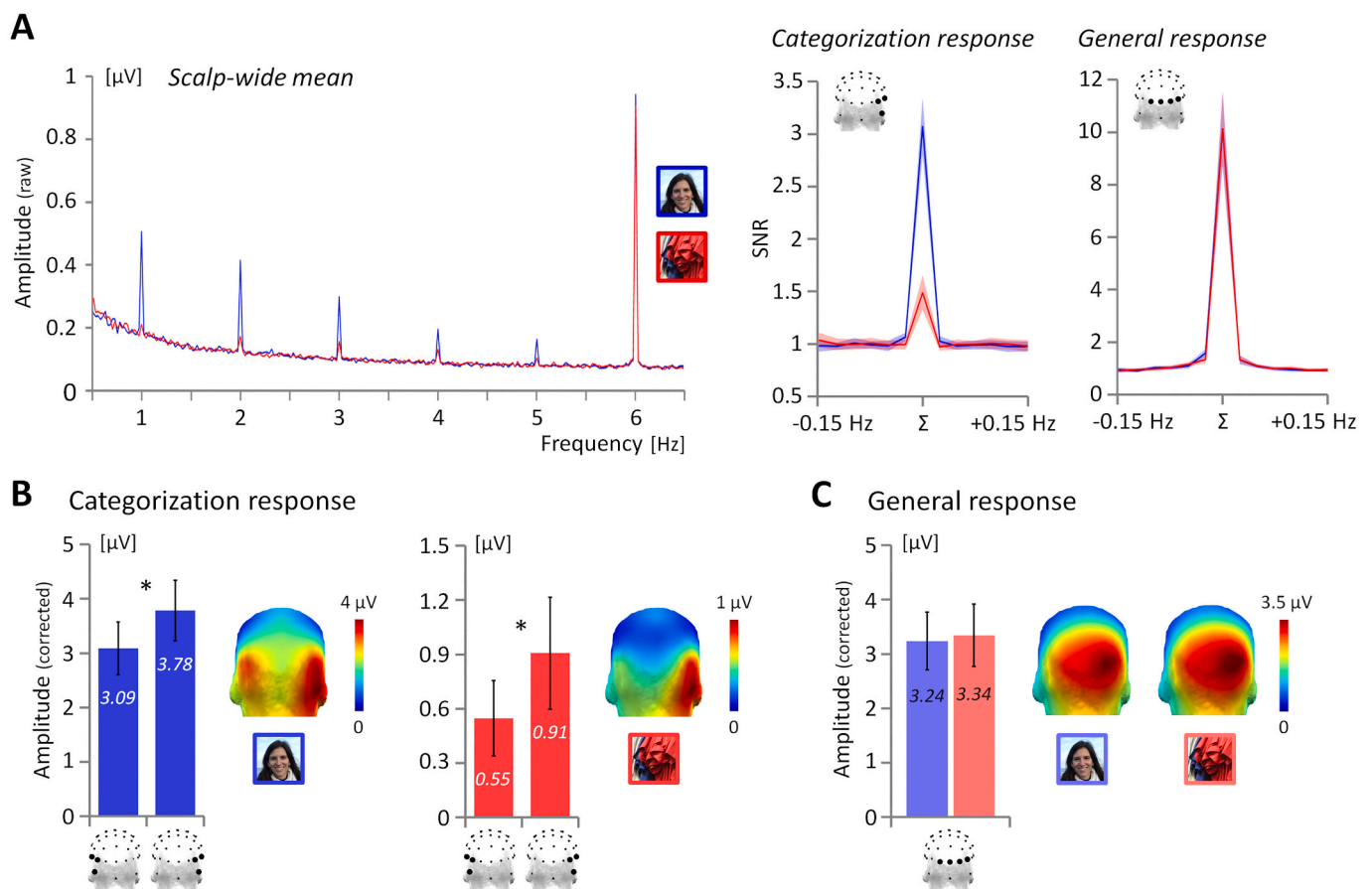
In a last step, we determined whether the facelike categorization response emerges as a function of participants' report of illusory faces for the combination of the two shortest durations (i.e., 17 ms and 33 ms) for which behavioral responses are not at ceiling. For each participant, we averaged preprocessed epochs (i.e., in the time domain) across durations separately for reported (i.e. *aware*) and unreported (i.e., *unaware*) sequences. Following processing steps were identical as in the

main analysis to obtain one brain response for perceptually aware sequences, and one response for unaware sequences. Individual summed corrected amplitudes were submitted to a repeated-measures ANOVA with *Awareness* (aware vs. unaware) and *Hemisphere* (rOT vs. lOT) as within-subject factors. Significance of the grand-averaged brain responses was estimated by determining whether the 95% CI did not include zero. For illustration, the difference between aware and unaware sequences was also computed for each participant. Finally, we conducted the same analysis with 12 participants out of 22 to control for a potential influence of the number of sequences associated with awareness or not, which is not equivalent across durations (see Fig. S2).

### 3. Results

#### 3.1. Experiment 1: characterizing the conscious categorization of illusory faces

At 1 Hz and harmonics, we identified two brain responses reflecting the categorization of human faces and facelike objects from variable natural images (Fig. 2A). Summed across harmonics, both categorization responses are of high signal-to-noise ratio (SNR  $\approx 3$  and 1.5



**Fig. 2.** Brain responses elicited in sequences presenting human faces (blue) or facelike objects (red) among nonface objects. **A.** Left: Grand-averaged FFT raw amplitude spectra (across 64 channels). Both types of stimulation sequences elicit a large response at the 6-Hz image presentation frequency (i.e., general response to all images). Though stronger to human faces, responses are clearly visible (i.e., of higher amplitude than surrounding frequency bins) at the 1-Hz human/illusory face presentation frequency (i.e., categorization response) and its harmonics (i.e., integer multiples, here 2 Hz, 3 Hz, 4 Hz, 5 Hz). Right: Grand-averaged signal-to-noise ratio (SNR) of the categorization and general responses over the right occipito-temporal and middle occipital regions, respectively. Responses are summed across significant harmonics ( $\Sigma$ ) and compared to surrounding frequencies ( $\pm 0.15$  Hz, SNR  $\approx 1$ , signal  $\approx$  noise). SNR is high for both responses (categorization response: SNR  $\approx 3$  and 1.5 for human faces and facelike objects, respectively; general response: SNR  $\approx 10$  for both categories). **B.** Grand-averaged summed corrected amplitude of the categorization responses over the left and right occipito-temporal regions ( $* p < .05$ ). **C.** Grand-averaged summed corrected amplitude of the general response over the middle occipital region. For both categorization (B) and general (C) responses, topographies are illustrated by head maps (posterior view). Shaded areas in (A) and error bars in (B) and (C) represent 95% confidence intervals. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)



respectively for human faces and facelike objects; i.e., 200% and 50% of signal increase) and significantly above noise level (i.e., 95% confidence intervals (CI) do not include 0; Fig. 2B), despite a larger response to human faces ( $3.44 \pm 0.45$  (95% CI)  $\mu\text{V}$ ) than facelike objects ( $0.72 \pm 0.19 \mu\text{V}$ ; 21% of the face categorization response;  $F_{1,26} = 175, p < .001, \eta_p^2 = 0.87$ ). Importantly, the two categorization responses present close topographies over the occipito-temporal cortex and a right hemisphere advantage (main effect of Hemisphere:  $F_{1,26} = 7.77, p = .009, \eta_p^2 = 0.23$ ). The human face categorization response is about 23% larger over the right (rOT;  $3.78 \pm 0.55 \mu\text{V}$ ) than the left occipito-temporal region (lOT;  $3.09 \pm 0.48 \mu\text{V}$ ) while the facelike categorization response is about 66% larger over rOT ( $0.91 \pm 0.31 \mu\text{V}$ ) than lOT ( $0.55 \pm 0.21 \mu\text{V}$ ). Considering the categorization responses over the whole scalp, about 11.6% and 16.4% of the human face and facelike categorization responses, respectively, are concentrated over rOT (representing less than 5% of the overall channels, i.e., 3 out of 64).

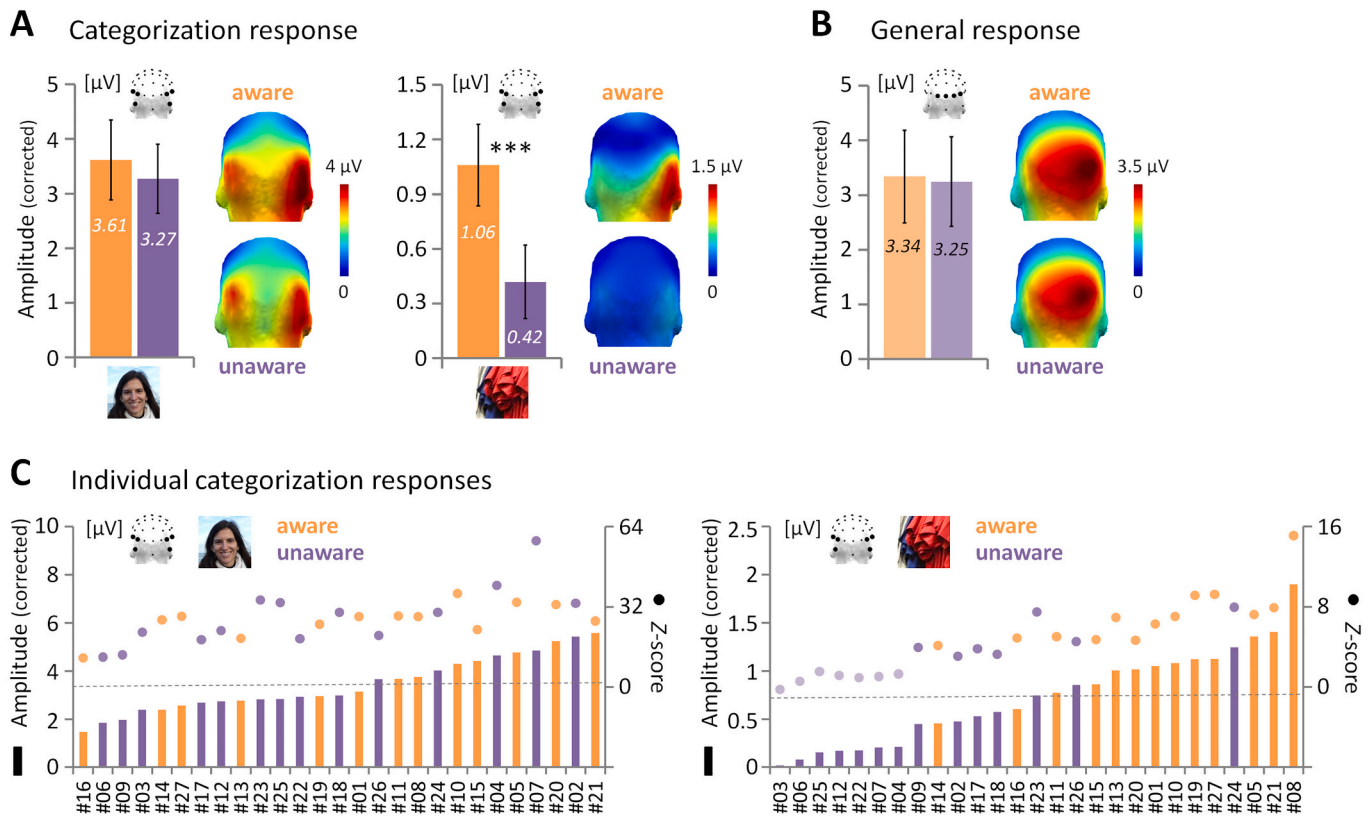
Using channels as observations, we confirmed the similar topographies of the two categorization responses (i.e., highly correlated;  $R = 0.92$ ). In comparison, the correlation between the facelike categorization response and the more central general visual response ( $R = 0.64$ ) is significantly lower ( $p < .001$ ). By splitting EEG data according to stimulus sets (see section 2.1.6.) and computing correlations between the responses obtained for each set, we also observed that both categorization responses to human faces ( $R = 0.99$ ) and facelike objects ( $R = 0.91$ ) are highly reliable across measurements.

It is worth noting that the visual categorization responses to human and illusory faces are automatically elicited since participants did not

explicitly process the two categories but performed an orthogonal cross-detection task with high accuracy ( $99.3 \pm 0.5\%$ ) and speed ( $406 \pm 11$  ms) without any difference between categories (both  $F_s < 1$ ). Similarly, the general visual response recorded at 6 Hz and harmonics over the middle occipital cortex (Fig. 2A & 2C) is not modulated by the stimulus category presented at 1 Hz (human faces:  $3.24 \pm 0.53 \mu\text{V}$ ; facelike objects:  $3.34 \pm 0.57 \mu\text{V}$ ;  $F < 1$ ).

Critically, after differentiating participants who overtly reported facelike objects (perceptually aware participants,  $N = 13$ ) and those who did not (unaware participants,  $N = 14$ ), the categorization response to facelike objects (Fig. 3A) is 153% larger for aware ( $1.06 \pm 0.22 \mu\text{V}$ ) than unaware participants ( $0.42 \pm 0.20 \mu\text{V}$ ;  $F_{1,25} = 21.3, p < .001, \eta_p^2 = 0.46$ ). In contrast, the human face categorization response is not significantly different between participants ( $F < 1$ ), albeit descriptively larger for aware ( $3.61 \pm 0.73 \mu\text{V}$ ; +10.5%) than unaware participants ( $3.27 \pm 0.63 \mu\text{V}$ ). Likewise, the amplitude of the general visual response (Fig. 3B) does not differ between participants (aware:  $3.34 \pm 0.85 \mu\text{V}$ ; unaware:  $3.25 \pm 0.82 \mu\text{V}$ ;  $F < 1$ ), who also do not differ at the behavioral cross-detection task (aware:  $99.5 \pm 0.4\%$ ,  $406 \pm 15$  ms; unaware:  $99.1 \pm 0.9\%$ ,  $406 \pm 17$  ms; both  $F_s < 1$ ).

Next, we estimated the significance of individual categorization responses (Fig. 3C) using Z-scores contrasting the amplitude of the response from surrounding noise level ( $Z > 1.64, p < .05$ , one-tailed, signal > noise). Every individual participant presents a strongly significant categorization response to human faces (all  $Z_s > 11.64$ , all  $p_s < 0.001$ ). In contrast, while the categorization response to illusory faces is significant in every perceptually aware participant (all  $Z_s > 4.15$ , all  $p_s$



**Fig. 3.** Categorization and general responses according to the perceptual awareness of illusory faces. A. Grand-averaged summed corrected amplitude of the categorization responses to human faces (left) and facelike objects (right) averaged across left and right occipito-temporal regions for the perceptually aware (orange) and unaware (purple) participants (\*\* $p < .001$ ). B. Grand-averaged summed corrected amplitude of the general response over the middle occipital region for the perceptually aware (orange) and unaware (purple) participants (categories collapsed). For both categorization (A) and general (B) responses, topographies are illustrated by head maps (posterior view) and error bars represent 95% confidence intervals. C. Individual categorization responses to human (left) and illusory (right) faces for perceptually aware (orange) and unaware (purple) participants. Bars and dots respectively depict summed corrected amplitudes (ranked in ascending order) and Z-scores. Lighter dots represent non-significant Z-scores ( $Z < 1.64, p > .05$ , one-tailed, signal > noise). The dashed grey line depicts the mean corrected amplitude across all participants. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

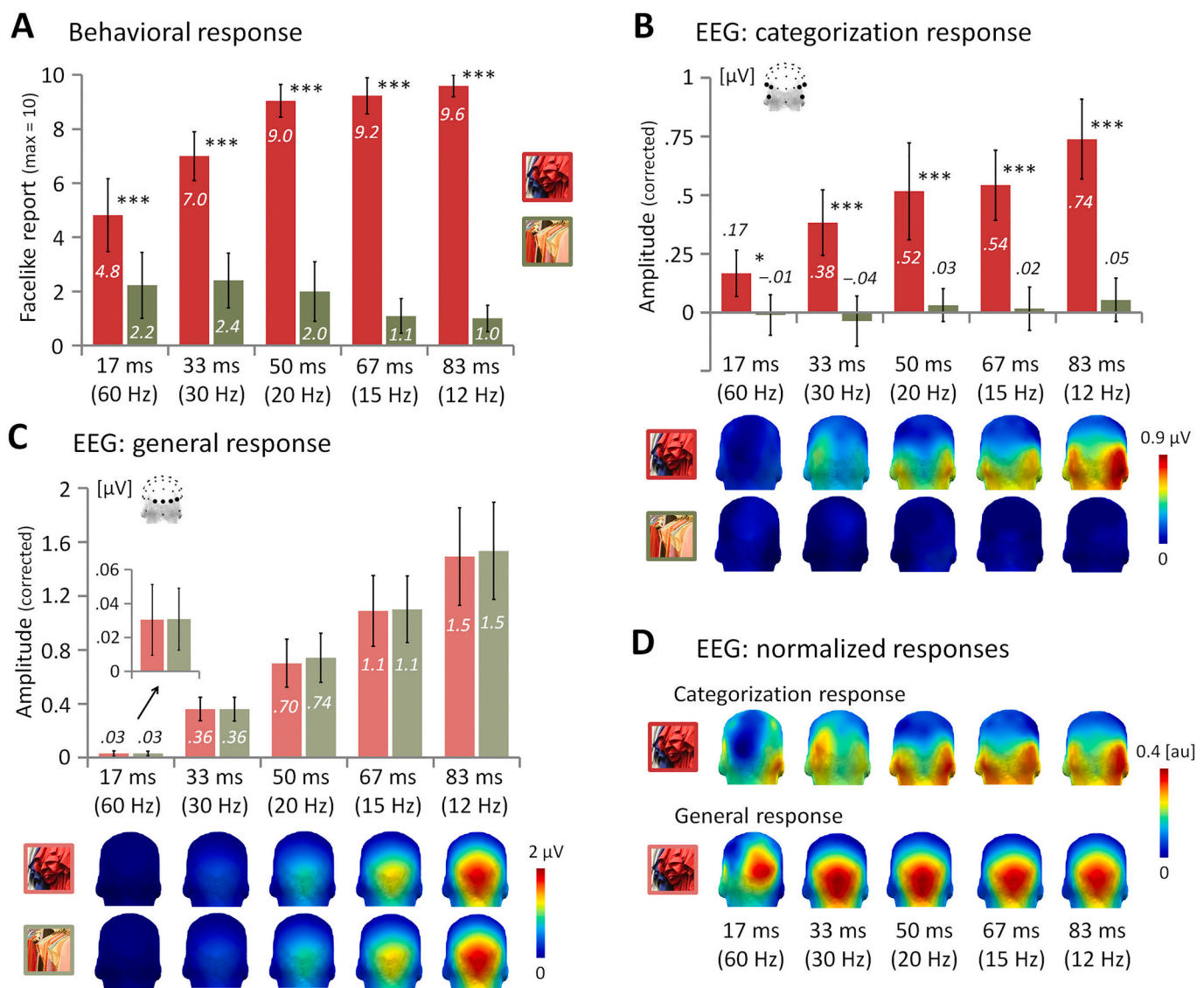
< 0.001), it is significant in only half of the unaware participants (i.e., 7 out of 14, all  $Z_s > 3.07$ , all  $p_s < 0.002$ , other half: from  $Z = -0.24$ ,  $p = .59$  to  $Z = 1.53$ ,  $p = .063$ ). By ranking the amplitude of individual categorization responses, we observed that 11 out of the 13 largest responses to facelike objects (i.e., above both the mean and the median responses) belong to perceptually aware participants. In other words, with this criterion, EEG data predicts well above chance ( $p = .011$ ) whether a given participant consciously perceives illusory faces (accuracy = 85%). In contrast, predictability is not above chance ( $p = .29$ ) if based on the response to human faces (8 out of the 13 largest responses (62%) belong to perceptually aware participants).

**3.2. Experiment 2: perceived vs. unperceived illusory faces in a single brain**

Experiment 1 demonstrates that variable objects resembling faces are categorized as faces by the human brain, in association with participants' report of face pareidolia. However, given that participants

were differentiated a posteriori from this single report, we conducted Experiment 2 to directly manipulate awareness in a within-subject design and compare the brain responses to perceived and unperceived facelike objects in each participant. After being presented with examples of facelike stimuli, another 22 participants were explicitly instructed to report if they perceived illusory faces after each of a hundred 16-s-long sequences. Half of the sequences presented facelike objects among nonface objects and the other half presented only nonface objects. Facelike stimuli were always displayed at 1 Hz, but they lasted 17 ms, 33 ms, 50 ms, 67 ms or 83 ms depending on the sequence (i.e., 5 stimulation frequencies: 60 Hz, 30 Hz, 20 Hz, 15 Hz and 12 Hz; Fig. 1C) to make the conscious perception of illusory faces challenging and dissociate sequences associated with perceptual awareness or not.

For sequences containing facelike objects (Fig. 4A), the mean number of facelike reports is low at 17 ms ( $4.8 \pm 1.4$  (95% CI) reports out of 10) and then increases at 33 ms ( $7.0 \pm 0.9$  reports) to reach near-ceiling accuracy from 50 ms ( $9.0 \pm 0.6$  reports) to 83 ms ( $9.6 \pm 0.4$  reports) with no difference between the three longest durations (all  $p_s > 0.05$ ). In



**Fig. 4.** Behavioral and neural responses according to stimulus category and duration. Mean number of facelike reports out of 10 (A), grand-averaged summed corrected amplitude of the categorization response over occipito-temporal regions (B) and of the general response over the middle occipital region (C) for sequences presenting facelike objects (red) or only nonface objects (green) and for the five stimulus durations (\*  $p < .05$ , \*\*\*  $p < .001$ , error bars represent 95% confidence intervals). For both categorization (B) and general (C) responses, topographies are illustrated by head maps (posterior view). D. Topographical head maps of normalized categorization and general responses for sequences containing facelike objects illustrate their spatial distribution across stimulus durations (au: arbitrary unit). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

contrast, for sequences containing only nonface objects, facelike reports are very low at all durations (from  $2.2 \pm 1.2$  reports at 17 ms to  $1.0 \pm 0.5$  reports at 83 ms, significant difference between 50 ms and 67 ms,  $p = .044$ ). Hence, the number of accurate facelike reports (i.e., hits) is always greater than the number of erroneous perceptions (i.e., false alarms), even at the shortest 17-ms duration (main effect of *Category*:  $F_{1,21} = 238$ ,  $p < .001$ ,  $\eta_p^2 = 0.92$ ). In other words, despite the very high constraints put on the visual system at the highest stimulation rate, participants are able to tell apart the two types of sequences (i.e., with or without facelike objects), this difference between the number of hits and false alarms increasing as a function of stimulus duration (*Category*  $\times$  *Stimulus duration* interaction:  $F_{2,5,52.3} = 37.5$ ,  $\epsilon = 0.62$ ,  $p < .001$ ,  $\eta_p^2 = 0.64$ ).

At the neural level, the response measured at 1 Hz and harmonics over occipito-temporal regions (Fig. 4B) is also always larger for sequences containing facelike stimuli than only nonface objects (main effect of *Category*:  $F_{1,21} = 57.8$ ,  $p < .001$ ,  $\eta_p^2 = 0.73$ ). Albeit low at the 17-ms stimulus duration ( $0.17 \pm 0.09 \mu\text{V}$ ), the response to facelike objects is greater than noise level (i.e., 95% CI does not include 0) and increases with stimulus duration (until  $0.74 \pm 0.17 \mu\text{V}$  at 83 ms), while there is no response to nonface objects (mean amplitude across durations:  $0.01 \pm 0.05 \mu\text{V}$ ). The difference between the responses to facelike objects and nonface objects thereby increases with duration (*Category*  $\times$  *Stimulus duration* interaction:  $F_{4,84} = 6.64$ ,  $p < .001$ ,  $\eta_p^2 = 0.24$ ). A right hemisphere advantage is visible at most durations for the categorization response to illusory faces (Fig. 4D), but it does not reach significance ( $F < 1$ ). Like the facelike categorization response, the general visual response (Fig. 4C) is larger than noise at all durations (lowest amplitude:  $0.03 \pm 0.02 \mu\text{V}$  at 17 ms) and increases with stimulus duration (until  $1.5 \pm 0.36 \mu\text{V}$  at 83 ms,  $F_{1,4,29.7} = 57.5$ ,  $\epsilon = 0.35$ ,  $p < .001$ ,  $\eta_p^2 = 0.73$ ). However, this middle occipital activity (Fig. 4D) elicited by the rapid stream of stimuli is not different between sequences containing facelike and nonface objects ( $F_{1,21} = 1.59$ ,  $p = .22$ ).

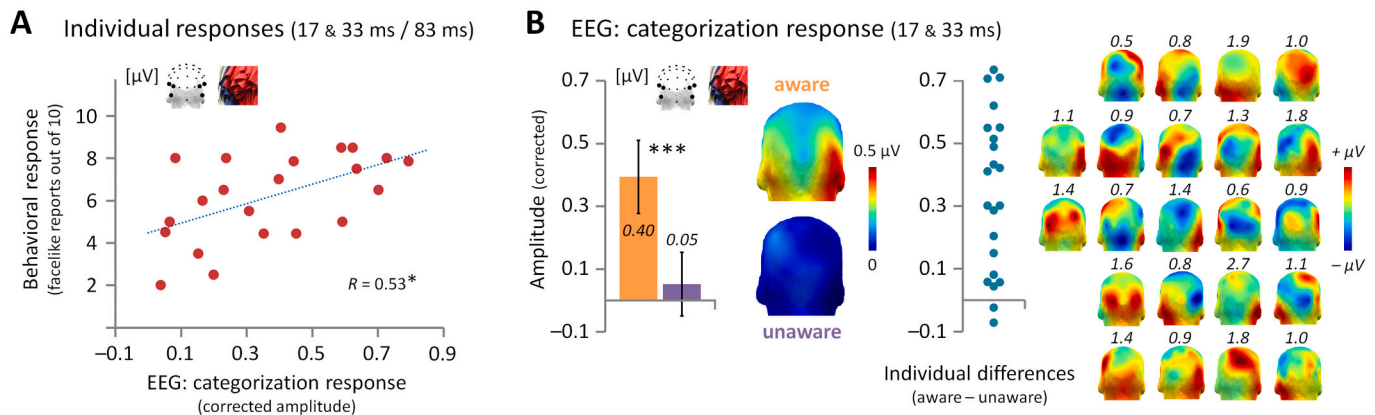
The increase of both neural and behavioral responses to facelike objects as stimulus duration increases suggests a relationship between these measures, in line with participants' awareness of face pareidolia. Thus, we evaluated whether participants presenting with larger facelike categorization responses report facelike stimuli in more sequences after weighting both responses by their value at the longest 83-ms duration to correct for ceiling-level neural activity and behavioral performance. For each duration individually, marginal relationships are observed only at

17 ms ( $R = 0.39$ ,  $p = .078$ ) and 33 ms ( $R = 0.42$ ,  $p = .053$ ; other durations: all  $R$ s  $< 0.22$ , all  $p$ s  $> 0.32$ ). When these two shortest durations are combined, neural and behavioral responses become significantly associated ( $R = 0.53$ ,  $p = .011$ ; Fig. 5A). In contrast, there is no association between the two measures for the combination of the two following durations (i.e., 50 and 67 ms;  $R = -0.01$ ,  $p = .97$ ), likely due to the near-ceiling accuracy observed for a majority of participants at these longer durations. In sum, for the two most challenging stimulus durations, which lead to more variable behavioral performance across participants, the overt report of illusory face perception is related to the amplitude of the facelike categorization response.

Accordingly, we finally explored whether the advent of the facelike categorization response at the 17- and 33-ms durations directly predicts participants' report of illusory faces. For each participant, we differentiated the facelike categorization responses between sequences wherein they perceived facelike objects (*aware* sequences) and those wherein they did not (*unaware* sequences). Strikingly, the categorization response to facelike stimuli is significantly above noise level only for perceptually aware sequences ( $0.40 \pm 0.12 \mu\text{V}$ ; Fig. 5B) and leads to a larger neural activity compared to unaware sequences ( $0.04 \pm 0.10 \mu\text{V}$ ;  $F_{1,21} = 40.2$ ,  $p < .001$ ,  $\eta_p^2 = 0.66$ ). Descriptively, 20 out of the 22 participants have a larger categorization response when they report facelike objects, such that the sign of the difference between the two conditions predicts above chance whether a given participant was aware of the illusory faces with an accuracy of 91% ( $p < .001$ ). The topography of this difference reveals an advantage for aware sequences over occipito-temporal scalp regions in every individual participant (Fig. 5B). These observations remain unchanged when only 12 participants are considered to equate the number of sequences associated with perceptual awareness or not across the two durations (Fig. S2). This demonstrates that the conscious perception of an illusory face emerges from neural facelike categorization.

#### 4. Discussion

Through two experiments, we identify a brain response to a variety of naturalistic facelike objects contrasted to many nonface objects of the same categories in a fast train of forward- and backward-masked stimuli, which, like the categorization response to genuine human faces, is mainly recorded over right occipito-temporal regions. Importantly, this valid neural measure of rapid and automatic facelike categorization



**Fig. 5.** The facelike categorization response predicts the conscious perception of an illusory face. A. Correlation between individual summed corrected amplitudes of the facelike categorization response over occipito-temporal regions and the number of facelike reports weighted by their ceiling-level value at the 83-ms duration for the combination of the 17- and 33-ms durations ( $* p < .05$ ). B. Left: grand-averaged summed corrected amplitude of the facelike categorization response over occipito-temporal regions depending on participants' report of illusory faces (perceptually aware sequences: orange, unaware sequences: purple;  $*** p < .001$ ) for the combination of the 17- and 33-ms durations. Error bars represent 95% confidence intervals. Topographies are illustrated by head maps (posterior view). Right: dots depict individual differences between the facelike categorization responses to reported (aware sequences) and unreported (unaware sequences) face pareidolia. Topographies are illustrated by head maps (posterior view) with each individual scale above the map. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)



under high visual constraints is isolated in individual participants and predicts perceptual awareness with high accuracy, either between groups of perceptually aware vs. unaware participants (Experiment 1), or between stimulation sequences according to participants' report of face pareidolia (Experiment 2). Hence, thanks to the advantages of EEG frequency-tagging to objectively measure automatic visual categorization in the brain, and capitalizing on a visual illusion to investigate the richness of face perception beyond the mere categorization of human faces, the present study provides a clear signature of face pareidolia in the human brain.

The first major achievement of the present study is to provide a rich and valid measure of facelike categorization under the variable viewing conditions in which it takes place in the natural visual environment, going beyond previous efforts with scalp EEG and neuroimaging (e.g., Caharel et al., 2013; Churches et al., 2009; Dolan et al., 1997; George et al., 2005; Rossion et al., 2011; but see Wardle et al., 2020). The use of various natural views of facelike objects contrasted to equivalent non-face stimuli makes unlikely the contribution of low physical variability between facelike objects, contrary to stimuli segmented from their background which artificially increases facelikeness by delineating a global face shape (see Davidenko et al., 2012 for a discussion). Naturalistic stimuli also implies figure-ground segregation, an integral part of object perception (Wagemans et al., 2012). In addition, the facelike categorization response reflects spontaneous face pareidolia, that is, the automatic perception (i.e., unintentional and hard to suppress) of an illusory face, at a glance. It is also worth noting that the facelike categorization response is a direct differential response (i.e., without post-hoc subtraction), such that it would have been absent if facelike objects were not discriminated from nonface objects, as observed in Experiment 2 for sequences wherein only nonface objects are displayed. The response is objective (i.e., recorded at a pre-experimentally defined frequency and its harmonics), and highly sensitive and reliable, as estimated in Experiment 1. These properties are critical to identify unambiguous individual brain responses, estimate their significance, and relate them to participants' report of face pareidolia.

In Experiment 1, we also clarify how similar is the categorization of illusory faces to the categorization of genuine human faces, both quantitatively and qualitatively. Consistent with previous studies using the same paradigm, we observe that the categorization of human faces elicits a large response of about 4  $\mu$ V over the occipito-temporal cortex, with a right hemisphere advantage (e.g., Jacques et al., 2016; Retter & Rossion, 2016). The facelike categorization response has a similar topography with an increased right-hemispheric dominance, but is of about 20% of the face categorization response in amplitude overall (35% when considering only perceptually aware participants). At least three non-mutually exclusive interpretations can explain this observation. First, the two responses may be generated by the same face-selective regions distributed along the VOTC (Gao et al., 2018; Jonas et al., 2016; Sergent et al., 1992; Zhen et al., 2015; Grill-Spector et al., 2017 for review), with a weaker activation overall for facelike objects. Such diminished responsiveness could be due to the absence of some cues pertaining to human faces in facelike stimuli, associated with a lower similarity across facelike objects than human faces and a higher similarity between facelike objects and nonface objects than between human faces and nonface objects (Wardle et al., 2020). For instance, while both shape and color information are important cues for visual recognition (Gegenfurtner & Rieger, 2000; Tanaka, Weiskopf, & Williams, 2001), and while color contributes to about 20% of the face categorization response for photographs (Or, Retter, & Rossion, 2019), color does not inform about facelikeness in facelike objects, is more variable across facelike than human face exemplars, and is more similar between facelike objects and nonface objects than between human faces and nonface objects. A second interpretation may be that only a subset of face-selective regions generates the facelike categorization response. Neuroimaging studies have found activity within the lateral part of the middle fusiform gyrus in response to facelike stimuli (Andrews &

Schluppeck, 2004; Dolan et al., 1997; Kanwisher et al., 1998; McKeeff & Tong, 2007; Rossion et al., 2011; Wardle et al., 2020), sometimes considered as a key region for conscious face perception (e.g., Grill-Spector, Knouf, & Kanwisher, 2004; Rodríguez et al., 2012), especially for the perception of a global face configuration (Andrews, Davies-Thompson, Kingstone, & Young, 2010; Goffaux, Schiltz, Mur, & Goebel, 2013; Rossion et al., 2011). Relatedly, the facelike categorization response is strongly right-lateralized, the right hemisphere being particularly involved in the perception of a global facelike configuration (Caharel et al., 2013; Parkin & Williamson, 1987; Rossion et al., 2011). However, despite a numerical advantage, there is no significant hemispheric asymmetry in Experiment 2, which may be due to the explicit instruction to attend to illusory faces increasing the left hemisphere contribution (see Quek, Nemrodov, Rossion, & Liu-Shuang, 2018). A last interpretation may be that facelike stimuli either elicit a full response within the whole face-selective network when they are perceived as faces, or they do not elicit a face-selective response at all when they are not perceived as faces, leading to a lower response in average. In other words, neural categorization could be strictly identical for facelike objects and human faces, but artificially appear weaker for facelike objects due to more occasional occurrences. Interestingly for our purpose, this account concurs with the view that face categorization emerges all at once from the linear accumulation of sensory evidence and reflects perceptual awareness (Harris et al., 2011; Navajas et al., 2013; Retter et al., 2020; Tong et al., 1998).

In that respect, the second major achievement of our study is to characterize to what extent the neural categorization of facelike objects reflects conscious illusory face perception, extending prior work on the association between a neural response to simple facelike stimuli and their perceptual interpretation as a face (Andrews & Schluppeck, 2004; Bentin et al., 2002; George et al., 2005; McKeeff & Tong, 2007; Shafto & Pitts, 2015). In Experiment 1, we observe a strong categorization response to facelike objects in participants who report face pareidolia post-stimulation compared to a weak response in those who do not, and reveal that individual facelike categorization responses predict this association. More strikingly, this relationship is confirmed in a single group of participants in Experiment 2, with a significant correlation between the facelike categorization response and the number of illusory face reports. As a result, when directly comparing sequences associated with awareness or not in this experiment, the facelike categorization response is observed only when face pareidolia are reported, with a conspicuous difference between the two types of sequences in each individual participant. Hence, given that the response varies greatly as a function of perceptual awareness despite identical stimuli, and given that these stimuli are directly contrasted to other stimuli depicting the same object categories, our study provides original evidence supporting the view that neural face categorization is a signature of conscious face perception (Harris et al., 2011; Navajas et al., 2013; Retter et al., 2020; Tong et al., 1998).

The strict absence of facelike categorization response to unreported face pareidolia in Experiment 2 points toward an all-or-none neural categorization function in response to sensory evidence gradually accumulating in early visual areas, as mentioned above. However, in Experiment 1, although the large difference between perceptually aware and unaware participants indicates that the bulk of the response reflects conscious illusory face perception, the response is not completely abolished in unaware participants. This suggests that a residual selective response to facelike objects could be observed in the absence of overt report. One explanation may be that some cues elicit a differential response between facelike and nonface objects, albeit non-sufficient to trigger full categorization. For instance, facelike objects all depict "eyelike" or "mouthlike" features (e.g., Fig. 1A), sometimes considered critical features to perceive a nonface stimulus as a face (e.g., Omer, Sapir, Hatuka, & Yovel, 2019). Similarly, some image statistics covary with facelikeness, such as more contrast (i.e., higher spatial frequencies) in the upper part of the image. This visual property is well-known to



already attract attention in newborns, as a precursor to develop face perception (e.g., Johnson, Dziurawiec, Ellis, & Morton, 1991; Simion & Di Giorgio, 2015 for review). Moreover, the presence of human faces in this experiment could have primed face-related cues in facelike stimuli. Alternatively, since perceptually aware and unaware participants in Experiment 1 were differentiated a posteriori following a single awareness assessment that summarizes the experience of almost 3000 stimulation cycles including more than 250 facelike stimuli, the small response in unaware participants could be due to an idiosyncratic confound such as the criteria to define an illusory face or the ability to remember sparse occurrences of face pareidolia. Therefore, the number of participants who consciously perceived illusory faces may have been underestimated in this experiment. In this context, it should be noted that both the general visual response elicited by the rapid stream of stimulation and the efficiency at the cross-detection task do not differ between participants, making unlikely the contribution of visual attention. Irrespective of these potential limitations, they do not concern Experiment 2, which reveals a striking difference between perceptually aware and unaware responses to facelike objects in a single group of participants.

By manipulating stimulus duration while participants explicitly process facelike stimuli, Experiment 2 additionally provides important information about the optimal conditions for face pareidolia to arise within a fast train of forward- and backward-masked stimuli. Albeit low at the shortest 17-ms duration, both facelike reports and the categorization response are higher for sequences containing facelike objects than only nonface objects at every duration. This means that participants are already able to perceive some illusory faces at 17 ms, in agreement with a minimal duration of approximately 13–17 ms to behaviorally or neurally categorize human faces in rapid streams of natural images (Keysers, Xiao, Földiák, & Perrett, 2001; Retter et al., 2020), or other visual objects in various experimental designs (Bacon-Macé, Macé, Fabre-Thorpe, & Thorpe, 2005; Fisch et al., 2009; Mohsenzadeh, Qin, Cichy, & Pantazis, 2018; Potter, Wyble, Haggmann, & McCourt, 2014). Facelike reports then increase at 33 ms and reach ceiling at 50 ms. In contrast, the amplitude of the facelike categorization response increases until the longest 83-ms duration and could still increase beyond (e.g., a larger response is observed with a 167-ms duration for perceptually aware participants in Experiment 1), unlike the response to human faces, which saturates at 83 ms (Retter et al., 2020). Moreover, behavioral and neural responses become decorrelated at 50 ms and 67 ms. This dissociation between behavioral and neural responses from 50 ms is not surprising given that one behavioral response was recorded after each sequence presenting at least 192 stimuli and sometimes including 16 facelike objects. Therefore, participants could have reported illusory face perception from a few facelike stimuli, such that accuracy rapidly reached ceiling at the intermediate 50-ms duration, but the number of categorized stimuli within a sequence could still increase beyond 50 ms. This accords with previous studies showing that various factors, such as the overlap of sensory information with forward and backward stimuli, lead to the categorization of only a fraction of stimuli at short presentation times (e.g., Bacon-Macé et al., 2005).

Interestingly, the number of false alarms only decreases from 67 ms, and is significantly higher than zero at all durations. Thus, the explicit instruction to detect facelike stimuli among other stimuli within rapid sequences depicting many object categories makes participants incline to falsely report facelike objects, even at ceiling-level stimulus durations. Importantly, however, false alarms are not associated with a significant neural categorization response. Since erroneous perceptions can be driven by any stimulus within a sequence of nonface stimuli, they do not occur periodically and do not translate at 1 Hz and harmonics in the EEG spectrum. Therefore, it should be further determined whether a facelike categorization response can be observed for any stimulus reported as facelike in a fast train of natural images by measuring a behavioral report for individual stimuli. Previous studies found that reported face perception in random patterns, such as pure-noise stimuli,

is associated with a facelike EEG response (Hansen, Thompson, Hess, & Elleberg, 2010; Smith, Gosselin, & Schyns, 2012) or neural activity within face-selective brain regions (Liu et al., 2014; Nestor, Vettel, & Tarr, 2013; Zhang et al., 2008). Together with the present observation that unreported facelike objects do not elicit a facelike categorization response, as previously observed for human faces (Retter et al., 2020), this clearly indicates that neural face(like) categorization reflects the subjective interpretation of a stimulus as a face, i.e., conscious face perception. However, this does not mean that face perception is purely top-down, which could be misinterpreted as “arising by chance for any stimulus”. For instance, face perception in pure-noise stimuli is not strictly random, as revealed by the classification images produced by merging the stimuli categorized as faces (Hansen et al., 2010; Liu et al., 2014; Nestor et al., 2013; Smith et al., 2012). It is visible in these images that noise configuration is coarsely facelike. Hence, even if neural face categorization reflects the subjective interpretation of a stimulus as a face, the brain always relies on a more or less well-defined “face signal” in that stimulus.

## 5. Conclusions

By using a widely variable set of naturalistic facelike objects contrasted to another variable set of the same object categories, we measure a rich neural categorization response to the facelike stimuli that is intimately related to their conscious perception as faces. Hence, coupling face pareidolia, which reflects the high inclination of the human visual system to perceive faces beyond genuine faces, and EEG frequency-tagging, which measures rapid categorization in the brain with objectivity, sensitivity, reliability and validity, provides a powerful approach to characterize the neural underpinnings of conscious face perception. In doing so, we corroborate the view that the perceptual awareness of a face emerges from a categorical response to unconscious sensory inputs (Harris et al., 2011; Navajas et al., 2013; Retter et al., 2020; Tong et al., 1998). More generally, we confirm the usefulness of visual illusions (e.g., “bistable” stimuli) to dissociate the subjective interpretation of a stimulus from its objective content (e.g., Blake & Logothetis, 2002; Brascamp, Sterzer, Blake, & Knapen, 2018; Gregory, 1997), as a proxy to measure perceptual awareness in the brain (Andrews, Schluppeck, Homfray, Matthews, & Blakemore, 2002; Brouwer, van Ee, & Schwarzbach, 2005; Srinivasan, Russell, Edelman, & TONI, 1999). In addition, while the present study is restricted to face (like) categorization and conscious face perception, the same relationship may likely generalize to other visual categories (Grill-Spector, Kushnir, Hendler, & Malach, 2000; Philastides & Sajda, 2006). Since visual categorization is subtended by a set of category-selective regions in the VOTC, as shown with the present paradigm in fMRI (Gao et al., 2018) and human intracerebral recordings (Hagen et al., 2020; Jonas et al., 2016), it is tempting to consider these regions as neural substrates of perceptual awareness, a long-standing issue in cognitive neuroscience (Boly et al., 2017; Dehaene, Lau, & Kouider, 2017; Koch, Massimini, Boly, & Tononi, 2016; Odegaard, Knight, & Lau, 2017 for reviews). We acknowledge that much research must be carried out to further clarify this issue.

## Data availability

EEG datasets and behavioral data have been deposited in the Open Science Framework and are freely available at <http://doi.org/10.17605/OSF.IO/CTZY2>

## Author contributions

D.R., J.-Y.B., R.B., B.R., and A.L. designed research; D.R., and A.L. performed research; D.R. and A.L. analyzed data; D.R. and A.L. wrote the paper; D.R., J.-Y.B., R.B., B.R., and A.L. revised versions of the paper; and J.-Y.B., and A.L. obtained funding.

## Declaration of Competing Interest

The authors declare no conflicts of interest.

## Acknowledgments

The authors are grateful to Romain Patruix and Lucas Ronat for their help in data collection. This work was supported by grants from the “Conseil Régional Bourgogne Franche-Comté” (PARI grant), the FEDER (European Funding for Regional Economic Development), the French “Investissements d’Avenir” program (project ISITE-BFC, contract ANR-15-IDEX-0003), and the French National Research Agency (ANR, contract ANR-19-CE28-0009).

## Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.cognition.2022.105016>.

## References

- Andrews, T. J., Davies-Thompson, J., Kingstone, A., & Young, A. W. (2010). Internal and external features of the face are represented holistically in face-selective regions of visual cortex. *Journal of Neuroscience*, 30(9), 3544–3552. <https://doi.org/10.1523/JNEUROSCI.4863-09.2010>
- Andrews, T. J., & Schluppeck, D. (2004). Neural responses to Mooney images reveal a modular representation of faces in human visual cortex. *NeuroImage*, 21(1), 91–98. <https://doi.org/10.1016/j.neuroimage.2003.08.023>
- Andrews, T. J., Schluppeck, D., Homfray, D., Matthews, P., & Blakemore, C. (2002). Activity in the Fusiform Gyrus Predicts Conscious Perception of Rubin’s Vase–Face Illusion. *NeuroImage*, 17(2), 890–901. <https://doi.org/10.1006/nimg.2002.1243>
- Aru, J., Axmacher, N., Lam, A. T. A. D., Fell, J., Elger, C. E., Singer, W., & Melloni, L. (2012). Local category-specific gamma band responses in the visual cortex do not reflect conscious perception. *Journal of Neuroscience*, 32(43), 14909–14914. <https://doi.org/10.1523/JNEUROSCI.2051-12.2012>
- Bacon-Macé, N., Macé, M. J.-M., 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. <https://doi.org/10.1016/j.visres.2005.01.004>
- Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society, Series B, Methodological*, 289–300. <https://doi.org/10.1111/j.2517-6161.1995.tb02031.x>
- Bentin, S., Allison, T., Puce, A., Perez, E., & McCarthy, G. (1996). Electrophysiological studies of face perception in humans. *Journal of Cognitive Neuroscience*, 8(6), 551–565. <https://doi.org/10.1162/jocn.1996.8.6.551>
- Bentin, S., Sagiv, N., Mecklinger, A., Friederici, A., & von Cramon, Y. D. (2002). Priming visual face-processing mechanisms: Electrophysiological evidence. *Psychological Science*, 13(2), 190–193. <https://doi.org/10.1111/1467-9280.00435>
- Blake, R., & Logothetis, N. K. (2002). Visual competition. *Nature Reviews Neuroscience*, 3(1), 13–21. <https://doi.org/10.1038/nrn701>
- Boly, M., Massimini, M., Tsuchiya, N., Postle, B. R., Koch, C., & Tononi, G. (2017). Are the neural correlates of consciousness in the front or in the Back of the cerebral cortex? Clinical and neuroimaging evidence. *The Journal of Neuroscience*, 37(40), 9603–9613. <https://doi.org/10.1523/JNEUROSCI.3218-16.2017>
- Brascamp, J., Sterzer, P., Blake, R., & Knapen, T. (2018). Multistable perception and the role of the Frontoparietal cortex in perceptual inference. *Annual Review of Psychology*, 69(1), 77–103. <https://doi.org/10.1146/annurev-psych-010417-085944>
- Brouwer, G. J., van Ee, R., & Schwarzbach, J. (2005). Activation in visual cortex correlates with the awareness of stereoscopic depth. *Journal of Neuroscience*, 25(45), 10403–10413. <https://doi.org/10.1523/JNEUROSCI.2408-05.2005>
- Caharel, S., Leleu, A., Bernard, C., Viggiano, M.-P., Lalonde, R., & Rebaï, R. S. J. (2013). Early holistic face-like processing of Arcimboldo paintings in the right occipito-temporal cortex: Evidence from the N170 ERP component. *International Journal of Psychophysiology*, 90(2), 157–164. <https://doi.org/10.1016/j.ijpsycho.2013.06.024>
- Churches, O., Baron-Cohen, S., & Ring, H. (2009). Seeing face-like objects: An event-related potential study. *NeuroReport*, 20(14), 1290–1294. <https://doi.org/10.1097/WNR.0b013e3283305a65>
- Crouzet, S. M., Kirchner, H., & Thorpe, S. J. (2010). Fast saccades toward faces: Face detection in just 100 ms. *Journal of Vision*, 10(4), 1–17. <https://doi.org/10.1167/10.4.16>
- Davidenko, N., Remus, D. A., & Grill-Spector, K. (2012). Face-likeness and image variability drive responses in human face-selective ventral regions. *Human Brain Mapping*, 33(10), 2334–2349. <https://doi.org/10.1002/hbm.21367>
- Dehaene, S., Lau, H., & Kouider, S. (2017). What is consciousness, and could machines have it? *Science*, 358(6362), 486–492. <https://doi.org/10.1126/science.aan8871>
- Dolan, R. J., Fink, G. R., Rolls, E., Booth, M., Holmes, A., Frackowiak, R. S. J., & Friston, K. J. (1997). How the brain learns to see objects and faces in an impoverished context. *Nature*, 389(6651), 596–599. <https://doi.org/10.1038/39309>
- Fisch, L., Privman, E., Ramot, M., Harel, M., Nir, Y., Kipervasser, S., ... Malach, R. (2009). Neural “ignition”: Enhanced activation linked to perceptual awareness in human ventral stream visual cortex. *Neuron*, 64(4), 562–574. <https://doi.org/10.1016/j.neuron.2009.11.001>
- Gao, X., Gentile, F., & Rossion, B. (2018). Fast periodic stimulation (FPS): A highly effective approach in fMRI brain mapping. *Brain Structure and Function*, 223(5), 2433–2454. <https://doi.org/10.1007/s00429-018-1630-4>
- Gegenfurtner, K. R., & Rieger, J. (2000). Sensory and cognitive contributions of color to the recognition of natural scenes. *Current Biology*, 10(13), 805–808. [https://doi.org/10.1016/S0960-9822\(00\)00563-7](https://doi.org/10.1016/S0960-9822(00)00563-7)
- George, N., Jemel, B., Fiori, N., Chaby, L., & Renault, B. (2005). Electrophysiological correlates of facial decision: Insights from upright and upside-down Mooney-face perception. *Cognitive Brain Research*, 24(3), 663–673. <https://doi.org/10.1016/j.cogbrainres.2005.03.017>
- Goffaux, V., Schiltz, C., Mur, M., & Goebel, R. (2013). Local discriminability determines the strength of holistic processing for faces in the fusiform face area. *Frontiers in Psychology*, 3. <https://doi.org/10.3389/fpsyg.2012.00604>
- Gregory, R. L. (1997). Knowledge in perception and illusion. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 352(1358), 1121–1127. <https://doi.org/10.1098/rstb.1997.0095>
- Grill-Spector, K., Knouf, N., & Kanwisher, N. (2004). The fusiform face area subserves face perception, not generic within-category identification. *Nature Neuroscience*, 7(5), 555–562. <https://doi.org/10.1038/nn1224>
- 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. <https://doi.org/10.1038/77754>
- Grill-Spector, K., Weiner, K. S., Gomez, J., Stigliani, A., & Natu, V. S. (2018). The functional neuroanatomy of face perception: From brain measurements to deep neural networks. *Interface Focus*, 8(4), 20180013. <https://doi.org/10.1098/rsfs.2018.0013>
- Grill-Spector, K., Weiner, K. S., Kay, K., & Gomez, J. (2017). The functional neuroanatomy of human face perception. *Annual Review of Vision Science*, 3(1), 167–196. <https://doi.org/10.1146/annurev-vision-102016-061214>
- Hadjikhani, N., Kverger, K., Naik, P., & Ahlfors, S. P. (2009). Early (M170) activation of face-specific cortex by face-like objects. *NeuroReport*, 20(4), 403–407. <https://doi.org/10.1097/WNR.0b013e328325a8e1>
- Hagen, S., Jacques, C., Maillard, L., Colnat-Coulbois, S., Rossion, B., & Jonas, J. (2020). Spatially dissociated intracerebral maps for face- and house-selective activity in the human ventral Occipito-temporal cortex. *Cerebral Cortex*, 30(7), 4026–4043. <https://doi.org/10.1093/cercor/bbaa022>
- Hansen, B. C., Thompson, B., Hess, R. F., & Ellemberg, D. (2010). Extracting the internal representation of faces from human brain activity: An analogue to reverse correlation. *NeuroImage*, 51(1), 373–390. <https://doi.org/10.1016/j.neuroimage.2010.02.021>
- 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(7:3), 1–12. <https://doi.org/10.1167/11.7.3>
- Hershler, O., Golan, T., Bentin, S., & Hochstein, S. (2010). The wide window of face detection. *Journal of Vision*, 10(10:21), 1–14. <https://doi.org/10.1167/10.10.21>
- Jacques, C., Retter, T. L., & Rossion, B. (2016). A single glance at natural face images generate larger and qualitatively different category-selective spatio-temporal signatures than other ecologically-relevant categories in the human brain. *NeuroImage*, 137, 21–33. <https://doi.org/10.1016/j.neuroimage.2016.04.045>
- Jeffreys, D. A. (1996). Evoked potential studies of face and object processing. *Visual Cognition*, 3(1), 1–38. <https://doi.org/10.1080/713756729>
- Johnson, M. H., Dzurawiec, S., Ellis, H., & Morton, J. (1991). Newborns’ preferential tracking of face-like stimuli and its subsequent decline. *Cognition*, 40(1), 1–19. [https://doi.org/10.1016/0010-0277\(91\)90045-6](https://doi.org/10.1016/0010-0277(91)90045-6)
- 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. *Proceedings of the National Academy of Sciences*, 113(28), E4088–E4097. <https://doi.org/10.1073/pnas.1522033113>
- Kanwisher, N., Tong, F., & Nakayama, K. (1998). The effect of face inversion on the human fusiform face area. *Cognition*, 68(1), B1–B11. [https://doi.org/10.1016/S0010-0277\(98\)00035-3](https://doi.org/10.1016/S0010-0277(98)00035-3)
- Keys, R. T., Taubert, J., & Wardle, S. G. (2021). A visual search advantage for illusory faces in objects. *Attention, Perception, & Psychophysics*, 83(5), 1942–1953. <https://doi.org/10.3758/s13414-021-02267-4>
- Keyser, C., Xiao, D.-K., Földiák, P., & Perrett, D. I. (2001). The speed of sight. *Journal of Cognitive Neuroscience*, 13(1), 90–101. <https://doi.org/10.1162/089892901564199>
- Koch, C., Massimini, M., Boly, M., & Tononi, G. (2016). Neural correlates of consciousness: Progress and problems. *Nature Reviews Neuroscience*, 17(5), 307–321. <https://doi.org/10.1038/nrn.2016.22>
- Liu, J., Li, J., Feng, L., Li, L., Tian, J., & Lee, K. (2014). Seeing Jesus in toast: Neural and behavioral correlates of face pareidolia. *Cortex*, 53, 60–77. <https://doi.org/10.1016/j.cortex.2014.01.013>
- Makeig, S., Bell, A. J., Jung, T.-P., & Sejnowski, T. J. (1996). Independent component analysis of electroencephalographic data. In D. Touretzky, M. Mozer, & M. Hasselmo (Eds.), *Vol. 8. Advances in neural information processing systems* (pp. 145–151). MIT Press.
- 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(3), 203–208. [https://doi.org/10.1016/0168-5597\(85\)90015-2](https://doi.org/10.1016/0168-5597(85)90015-2)
- McKeeff, T. J., & Tong, F. (2007). The timing of perceptual decisions for ambiguous face stimuli in the human ventral visual cortex. *Cerebral Cortex*, 17(3), 669–678. <https://doi.org/10.1093/cercor/bkh015>

- Merriam-Webster. (2021). *Pareidolia*. Retrieved November 2, 2021, from <https://www.merriam-webster.com/dictionary/pareidolia>.
- 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, Article e36329. <https://doi.org/10.7554/eLife.36329>
- Moulson, M. C., Balas, B., Nelson, C., & Sinha, P. (2011). EEG correlates of categorical and graded face perception. *Neuropsychologia*, 49(14), 3847–3853. <https://doi.org/10.1016/j.neuropsychologia.2011.09.046>
- Moutoussis, K., & Zeki, S. (2002). The relationship between cortical activation and perception investigated with invisible stimuli. *Proceedings of the National Academy of Sciences*, 99(14), 9527–9532. <https://doi.org/10.1073/pnas.142305699>
- Navajas, J., Ahmadi, M., & Quiroga, R. Q. (2013). Uncovering the mechanisms of conscious face perception: A single-trial study of the N170 responses. *Journal of Neuroscience*, 33(4), 1337–1343. <https://doi.org/10.1523/JNEUROSCI.1226-12.2013>
- Nestor, A., Vettel, J. M., & Tarr, M. J. (2013). Internal representations for face detection: An application of noise-based image classification to BOLD responses. *Human Brain Mapping*, 34(11), 3101–3115. <https://doi.org/10.1002/hbm.22128>
- Norcia, A. M., Appelbaum, L. G., Ales, J. M., Cottareau, B. R., & Rossion, B. (2015). The steady-state visual evoked potential in vision research: A review. *Journal of Vision*, 15(6:4), 1–46. <https://doi.org/10.1167/15.6.4>
- Odegaard, B., Knight, R. T., & Lau, H. (2017). Should a few null findings falsify prefrontal theories of conscious perception? *Journal of Neuroscience*, 37(40), 9593–9602. <https://doi.org/10.1523/JNEUROSCI.3217-16.2017>
- Omer, Y., Sapir, R., Hatuka, Y., & Yovel, G. (2019). What is a face? Critical features for face detection. *Perception*, 48(5), 437–446. <https://doi.org/10.1177/0301006619838734>
- Or, C. C.-F., Retter, T. L., & Rossion, B. (2019). The contribution of color information to rapid face categorization in natural scenes. *Journal of Vision*, 19(5:20), 1–20. <https://doi.org/10.1167/19.5.20>
- Parkin, A. J., & Williamson, P. (1987). Cerebral lateralisation at different stages of facial processing. *Cortex*, 23(1), 99–110. [https://doi.org/10.1016/S0010-9452\(87\)80022-9](https://doi.org/10.1016/S0010-9452(87)80022-9)
- Perry, G. (2016). The visual gamma response to faces reflects the presence of sensory evidence and not awareness of the stimulus. *Royal Society Open Science*, 3(3), Article 150593. <https://doi.org/10.1098/rsos.150593>
- Philiastides, M. G., & Sajda, P. (2006). Temporal characterization of the neural correlates of perceptual decision making in the human brain. *Cerebral Cortex*, 16(4), 509–518. <https://doi.org/10.1093/cercor/bhi130>
- Potter, M. C., Wyble, B., Haggmann, C. E., & McCourt, E. S. (2014). Detecting meaning in RSVP at 13 ms per picture. *Attention, Perception, & Psychophysics*, 76(2), 270–279. <https://doi.org/10.3758/s13414-013-0605-z>
- Quek, G. L., Nemrodov, D., Rossion, B., & Liu-Shuang, J. (2018). Selective attention to faces in a rapid visual stream: Hemispheric differences in enhancement and suppression of category-selective neural activity. *Journal of Cognitive Neuroscience*, 30(3), 393–410. [https://doi.org/10.1162/jocn\\_a\\_01220](https://doi.org/10.1162/jocn_a_01220)
- Quek, G. L., & Rossion, B. (2017). Category-selective human brain processes elicited in fast periodic visual stimulation streams are immune to temporal predictability. *Neuropsychologia*, 104, 182–200. <https://doi.org/10.1016/j.neuropsychologia.2017.08.010>
- Retter, T. L., Jiang, F., Webster, M. A., & Rossion, B. (2020). All-or-none face categorization in the human brain. *NeuroImage*, 213, Article 116685. <https://doi.org/10.1016/j.neuroimage.2020.116685>
- 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. <https://doi.org/10.1016/j.neuropsychologia.2016.07.028>
- Robert, F., & Robert, J. (1996). *Face to face*. Lars Müller Publishers.
- Rodríguez, V., Thompson, R., Stokes, B., Brett, M., Alvarez, I., Valdes-Sosa, M., & Duncan, J. (2012). Absence of face-specific cortical activity in the complete absence of awareness: Converging evidence from functional magnetic resonance imaging and event-related potentials. *Journal of Cognitive Neuroscience*, 24(2), 396–415. [https://doi.org/10.1162/jocn\\_a\\_00137](https://doi.org/10.1162/jocn_a_00137)
- Rossion, B. (2014). Understanding face perception by means of human electrophysiology. *Trends in Cognitive Sciences*, 18(6), 310–318. <https://doi.org/10.1016/j.tics.2014.02.013>
- Rossion, B., Dricot, L., Goebel, R., & Busigny, T. (2011). Holistic face categorization in higher order visual areas of the Normal and Prosopagnosic brain: Toward a non-hierarchical view of face perception. *Frontiers in Human Neuroscience*, 4. <https://doi.org/10.3389/fnhum.2010.00225>
- Rossion, B., Torfs, K., Jacques, C., & Liu-Shuang, J. (2015). Fast periodic presentation of natural images reveals a robust face-selective electrophysiological response in the human brain. *Journal of Vision*, 15(1), 18. <https://doi.org/10.1167/15.1.18>
- Sagiv, N., & Bentin, S. (2001). Structural encoding of human and schematic faces: Holistic and part-based processes. *Journal of Cognitive Neuroscience*, 13(7), 937–951. <https://doi.org/10.1162/089992901753165854>
- Scheirer, W. J., Anthony, S. E., Nakayama, K., & Cox, D. D. (2014). Perceptual annotation: Measuring human vision to improve computer vision. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 36(8), 1679–1686. <https://doi.org/10.1109/TPAMI.2013.2297711>
- Sergent, J., Ohta, S., & Macdonald, B. (1992). Functional neuroanatomy of face and object processing: A positron emission tomography study. *Brain*, 115(1), 15–36. <https://doi.org/10.1093/brain/115.1.15>
- Shafto, J. P., & Pitts, M. A. (2015). Neural signatures of conscious face perception in an inattentive blindness paradigm. *Journal of Neuroscience*, 35(31), 10940–10948. <https://doi.org/10.1523/JNEUROSCI.0145-15.2015>
- Simion, F., & Di Giorgio, E. (2015). Face perception and processing in early infancy: Inborn predispositions and developmental changes. *Frontiers in Psychology*, 6. <https://doi.org/10.3389/fpsyg.2015.00969>
- Smith, M. L., Gosselin, F., & Schyns, P. G. (2012). Measuring internal representations from behavioral and brain data. *Current Biology*, 22(3), 191–196. <https://doi.org/10.1016/j.cub.2011.11.061>
- Srinivasan, R., Russell, D. P., Edelman, G. M., & Tononi, G. (1999). Increased synchronization of Neuromagnetic responses during conscious perception. *The Journal of Neuroscience*, 19(13), 5435–5448. <https://doi.org/10.1523/JNEUROSCI.19-13-05435.1999>
- Tanaka, J., Weiskopf, D., & Williams, P. (2001). The role of color in high-level vision. *Trends in Cognitive Sciences*, 5(5), 211–215. [https://doi.org/10.1016/S1364-6613\(00\)01626-0](https://doi.org/10.1016/S1364-6613(00)01626-0)
- Tanskanen, T., Näsänen, R., Ojanpää, H., & Hari, R. (2007). Face recognition and cortical responses: Effect of stimulus duration. *NeuroImage*, 35(4), 1636–1644. <https://doi.org/10.1016/j.neuroimage.2007.01.023>
- Taubert, J., Flessert, M., Wardle, S. G., Basile, B. M., Murphy, A. P., Murray, E. A., & Ungerleider, L. G. (2018). Amygdala lesions eliminate viewing preferences for faces in rhesus monkeys. *Proceedings of the National Academy of Sciences*, 115(31), 8043–8048. <https://doi.org/10.1073/pnas.1807245115>
- Taubert, J., Wardle, S. G., Flessert, M., Leopold, D. A., & Ungerleider, L. G. (2017). Face Pareidolia in the Rhesus monkey. *Current Biology*, 27(16), 2505–2509.e2. <https://doi.org/10.1016/j.cub.2017.06.075>
- Taubert, J., Wardle, S. G., & Ungerleider, L. G. (2020). What does a “face cell” want? *Progress in Neurobiology*, 195, Article 101880. <https://doi.org/10.1016/j.pneurobio.2020.101880>
- Tong, F., Nakayama, K., Vaughan, J. T., & Kanwisher, N. (1998). Binocular rivalry and visual awareness in human Extrastriate cortex. *Neuron*, 21(4), 753–759. [https://doi.org/10.1016/S0896-6273\(00\)80592-9](https://doi.org/10.1016/S0896-6273(00)80592-9)
- Vuong, Q. C., Willenbockel, V., Zimmermann, F. G. S., Lochy, A., Laguesse, R., Dryden, A., & Rossion, B. (2017). Facelikeness matters: A parametric multipart object set to understand the role of spatial configuration in visual recognition. *Visual Cognition*, 24(7–8), 406–421. <https://doi.org/10.1080/13506285.2017.1289997>
- Wagemans, J., Elder, J. H., Kubovy, M., Palmer, S. E., Peterson, M. A., Singh, M., & von der Heydt, R. (2012). A century of gestalt psychology in visual perception: I. perceptual grouping and figure-ground organization. *Psychological Bulletin*, 138(6), 1172–1217. <https://doi.org/10.1037/a0029333>
- Wardle, S. G., Taubert, J., Teichmann, L., & Baker, C. I. (2020). Rapid and dynamic processing of face pareidolia in the human brain. *Nature Communications*, 11(1), 4518. <https://doi.org/10.1038/s41467-020-18325-8>
- Zhang, H., Liu, J., Huber, D. E., Rieth, C. A., Tian, J., & Lee, K. (2008). Detecting faces in pure noise images: A functional MRI study on top-down perception. *NeuroReport*, 19(2), 229–233. <https://doi.org/10.1097/WNR.0b013e3282f49083>
- Zhen, Z., Yang, Z., Huang, L., Kong, X., Wang, X., Dang, X., Huang, Y., Song, Y., & Liu, J. (2015). Quantifying interindividual variability and asymmetry of face-selective regions: A probabilistic functional atlas. *NeuroImage*, 113, 13–25. <https://doi.org/10.1016/j.neuroimage.2015.03.010>