Cerebral Cortex, 2021;00: 1–14

https://doi.org/10.1093/cercor/bhab238
Advance Access Publication Date: 9 September 2021
Original Article

ORIGINAL ARTICLE

Coarse-to-Fine(r) Automatic Familiar Face Recognition in the Human Brain

Xiaoqian Yan[1,2,3], Valérie Goffaux[3,4,5] and Bruno Rossion[2,3,6]

[1]Department of Psychology, Stanford University, Palo Alto, CA 94305, USA, [2]Université de Lorraine, CNRS, CRAN, 54000 Nancy, France, [3]Institute of Research in Psychology (IPSY), University of Louvain, Louvain-La-Neuve 1348, Belgium, [4]Department of Cognitive Neuroscience, Maastricht University, Maastricht, 6229, the Netherlands, [5]Institute of Neuroscience (IoNS), University of Louvain, Louvain-La-Neuve 1348, Belgium and [6]Université de Lorraine, CHRU-Nancy, Service de Neurologie, 54000 Nancy, France

Address correspondence to Bruno Rossion, CRAN UMR 7039, CNRS—Universite de Lorraine, 2 Avenue de la forêt de Haye, 54516 Vandoeuvre-lès-Nancy, France. Email: bruno.rossion@univ-lorraine.fr

Abstract

At what level of spatial resolution can the human brain recognize a familiar face in a crowd of strangers? Does it depend on whether one approaches or rather moves back from the crowd? To answer these questions, 16 observers viewed different unsegmented images of unfamiliar faces alternating at 6 Hz, with spatial frequency (SF) content progressively increasing (i.e., coarse-to-fine) or decreasing (fine-to-coarse) in different sequences. Variable natural images of celebrity faces every sixth stimulus generated an objective neural index of single-glanced automatic familiar face recognition (FFR) at 1 Hz in participants' electroencephalogram (EEG). For blurry images increasing in spatial resolution, the neural FFR response over occipitotemporal regions emerged abruptly with additional cues at about 6.3–8.7 cycles/head width, immediately reaching amplitude saturation. When the same images progressively decreased in resolution, the FFR response disappeared already below 12 cycles/head width, thus providing no support for a predictive coding hypothesis. Overall, these observations indicate that rapid automatic recognition of heterogenous natural views of familiar faces is achieved from coarser visual inputs than generally thought, and support a coarse-to-fine FFR dynamics in the human brain.

Key words: coarse-to-fine, EEG, familiar face recognition, frequency-tagging, spatial frequency

Introduction

Recognizing a face as being familiar, that is, previously encoded as a distinctive face identity in memory, is fundamental for the quality of social interactions in the human species. Human face recognition has been scientifically studied for decades with 2D images (Ellis 1975; Bruce and Young 1986; Calder et al. 2011 for reviews). Here we use such images to ask 2 related questions regarding familiar face recognition (FFR). First: "how much spatial resolution does the human brain require to recognize a natural image of a face as being familiar?" We aim at answering this question for neurotypical human adults, who are generally considered as having the highest level of expertise at face identity recognition, at least compared with infants and children, various neurological and neuropsychiatric populations, and other animal species (Rossion 2018).

Given the importance of this issue for clinical (i.e., visual impairment) and applied research (i.e., the development and optimization of artificial face recognition systems), many studies have investigated the spatial resolution—either in terms of the number of pixels by image or more often in terms of the scale of luminance variations in an image (i.e., spatial frequencies, SF; Campbell, Cooper, and Enroth-Cugell 1969; De Valois and De Valois 1980)—that is necessary to recognize face identities. According to early reports/illustrations, face identities could be well recognized even at extremely low spatial resolution (Harmon 1973; Harmon and Julesz 1973; Ginsburg 1980; Rubin and Siegel 1984; Sinha 2002; see also Yip and Sinha 2002). Subsequent, more systematic, experimental studies found that face identity recognition drops significantly below about 6–8 SF cycles/face width (Bindemann et al. 2013; Fiorentini et al. 1983; Peli

et al. 1994) and does not increase significantly beyond this range (Bachmann 1991; Costen et al. 1994; Peli et al. 1994). Despite these observations, the prevalent view in the scientific community is that identity cues provided by a medium range of SF (e.g., between 8 and 12 or 8 and 16 cycles/face width) may be optimal for face identity recognition (i.e., "The golden mean", e.g., Collin et al. 2006; Collin et al. 2014; Costen, Parker, & Craw 1996; Fiorentini et al. 1983; Gao & Maurer 2011; Gold, Bennett, & Sekuler 1999; Keil 2008; Näsänen 1999; Ojanpää & Näsänen 2003; Parker & Costen 1999; Schyns, Bonnar, Gosselin 2002; Tanskanen et al. 2005; Watier et al. 2010).

The diversity of outcomes across behavioral studies that have investigated this issue could be largely attributed to substantial differences between experiments in terms of stimulation (e.g., number of individual faces, physical variability between faces, variable stimulus duration, different methods of spatial filtering through quantization, filters or band-pass noise added to images, etc.) and task requirements (matching pictures of the same identity against different distractors, identifying single pictures of familiar faces, categorizing familiar and unfamiliar faces, etc.). However, despite these methodological differences, previous studies also have in common 3 characteristics that may limit their conclusions. First, participants in these experiments are typically asked to recognize familiar or (more often) familiarized facial identities with only one picture per face, well segmented from its natural background and without diagnostic external cues (e.g., the hair), usually at full-front view only, among a visually homogenous set of faces. Yet, as particularly emphasized in human face recognition research in the last decade, using natural (i.e., unsegmented) images of faces that are widely variable in viewing conditions (e.g., lighting, head orientation, expression, etc.) for a given identity greatly increases the validity of facial recognition measures (e.g., Burton 2013; Young & Burton 2017). To date, the spatial resolution required to recognize face identities in such natural stimulus conditions remains unknown. Second, previous studies have used "explicit" recognition tasks varying substantially in terms of memory-related, attentional, and decisional processes, which may contribute greatly to the variability of findings across studies (Ruiz-Soler & Beltran 2006). However, FFR in neurotypical human adults is largely automatic (i.e., not under volitional control, e.g., Yan, Young, & Andrews 2017; Yan & Rossion 2020) and therefore does not require explicit tasks. Measuring FFR in an implicit manner would be particularly welcome to minimize variability due to task requirements and provide a stable measure of FFR under various spatial frequency contents. Finally, most SF studies with face stimuli presented those stimuli for a relatively long (or even unlimited) time, allowing observers to explore them with eye-movements that can introduce undesirable confounds (i.e., substantial variations of number of eye fixations and saccadic length with spatial frequency content; Ojanpää & Näsänen 2003 for faces; see also Groner et al. 2008; Tavassoli et al. 2009). This is particularly unfortunate because familiar faces can usually be recognized very rapidly, even at a single glance (Barragan-Jason et al. 2015; Caharel et al. 2014; di Oleggio Castello & Gobbini 2015; Hacker et al. 2019; Hsiao & Cottrell 2008; Yan & Rossion 2020), making long presentation durations unnecessary.

Here, taking into account all of these issues at once, we provide an implicit measure of human FFR as a function of spatial frequency with briefly presented natural and variable views of facial identities. To do so, we rely on a dynamic visual stimulation paradigm in which variable natural (i.e., unsegmented) images of unfamiliar faces alternate at a fixed rate of 6 Hz (i.e., 6 images by second) (Fig. 1), allowing human observers only a single glance at each face identity. Critically, highly variable images of different familiar (i.e., famous) face identities are embedded every second (i.e., 1 Hz) in the stimulation sequence (Fig. 1). Thus, in the high-density electroencephalogram (EEG) recordings of these participants, neural activity at 6 Hz reflects general visual processes common to all face stimuli, whereas activity at 1 Hz (and harmonics), if present, reflect a specific response to familiar faces (i.e., a neural measure of FFR) (Yan & Rossion 2020; see also Campbell et al. 2020; Yan, Zimmermann, & Rossion 2020; Zimmermann, Yan, & Rossion 2019). Importantly, given that many (i.e., 20) different natural views of each (familiar and unfamiliar) identity and 6 different familiar face identities are presented in a stimulation sequence, the FFR response obtained in the EEG is not tied to identity-related physical properties of these images (see Yan & Rossion 2020). Accordingly, and in line with large effects of picture-plane inversion observed in behavioral measures of face identity recognition or face familiarity judgments (e.g., Busigny & Rossion 2010; Collishaw & Hole 2000; Yin 1969), the FFR neural response is reduced of about 83% of amplitude when the exact same stimuli appear upside-down, preserving all physical differences between images (Yan & Rossion 2020).

Here we evaluate this neural measure of FFR in a sweep frequency-tagging paradigm (Regan 1973; see e.g., Ales et al. 2012; Quek et al. 2018 for face stimuli) by presenting spatially low-pass filtered images that progressively increase in SF content (i.e., coarse-to-fine presentation mode; CF; Fig. 2) in half of the stimulation sequences. Importantly, in the other half of stimulation sequences, the order of presentation is reversed, so that the highest resolution images progressively decrease in SF content, as if they were progressively low-pass filtered (i.e., fine-to-coarse, FC, Supplementary Fig. S1). The comparison between the results obtained in the CF and FC sequences (Bruner & Potter 1964) provides answer to the second question of interest of our study: "whether rapid and automatic FFR is affected by the temporal order of spatial information presentation, and if so how?" That is, given the same level of image spatial resolution, is there any recognition advantage when the familiar face identity has already—recently—been viewed at a higher spatial resolution (i.e., including more details)? On the one hand, a FC mode of stimulation, in which information is progressively lost, should provide observers with an advantage in recognizing familiar faces at a lower level of spatial resolution than in a CF mode, in which information progressively accumulates (i.e., a perceptual hysteresis effect in the temporal integration of high and low SF face information, see Bruner & Potter 1964; Brady & Oliva 2012). This hypothesis is in line with the influential predictive coding framework, according to which our prior experiences are constantly used to form predictions of upcoming events to ensure efficient processing (Friston 2005; Rao & Ballard 1999; see Trapp, Pascucci, & Chelazzi 2021 for a recent discussion of perceptual hysteresis and predictive coding). On the other hand, the visual system is thought to naturally accumulate information in a coarse-to-fine manner, that is, with lower resolution cues being extracted before finer details (Bachmann 1991; Goffaux et al. 2011; Hegdé 2008; Parker et al. 1992; Petras et al. 2019; Sergent 1986; Watt 1987). Hence, according to this latter view, recognition based on higher details—or high spatial frequency cues—in a face stimulus should rather be facilitated when lower resolution information has been presented before, as in a "natural" order. To contrast these predictions, in both stimulation modes, we measure the threshold, that is, the spatial frequency content level at which the first (in the CF condition)/last (in the FC condition)

A

SF STEP	1	2	3	4	5	6	7	8	9
cycles/image (cpi)	3	4.15	5.7	7.92	10.95	15.14	20.93	28.94	40
cycles/degree (cpd)	0.31	0.43	0.59	0.81	1.13	1.56	2.15	2.98	4.12
cycles/head width (cphw) (SD)	1.92 (0.23)	2.66 (0.31)	3.65 (0.43)	5.07 (0.6)	7.01 (0.82)	9.69 (1.14)	13.4 (1.57)	18.52 (2.17)	25.6 (3.01)

B

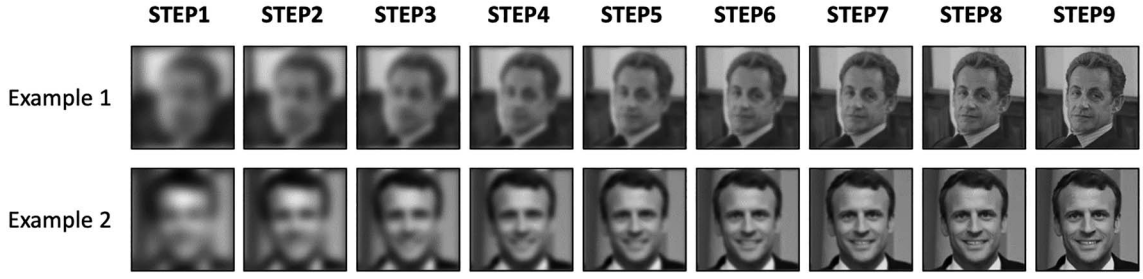

Figure 1. (*A*) The table indicates the low-pass SF cutoffs in terms of different units as used in the present study. Cycles/head width (cphw) was calculated based on the mean head width across all 120 familiar faces. Values in the parentheses indicate standard deviation of the mean. (*B*) Example face images filtered at each of the 9 logarithmic cut-off values. Face images shown here are with license permits. For license information, Nicolas Sarkozy, and Emmanuel Macron: Pictures under the Creative Commons Attribution 2.0 Generic. Attribution: Tallinn Digital and European People's Party, respectively.

significant neural FFR response is recorded, the saturation point (i.e., optimal level for FFR), and the evolution of this response, that is, whether it increases/decreases progressively or abruptly.

Materials and Methods

Participants

We recruited 18 participants in the experiment. The data of 2 participants were excluded due to excessive noise/muscular artifacts during EEG recording. The final sample consisted of 16 participants (7 females, mean age 22.5 ± 1.6 years). All participants reported to have normal/corrected-to-normal vision; none had a history of neurological or psychiatric disorder. They were all right-handed according to self-report. Written consent was obtained prior to test for a study approved by the Biomedical Ethical committee of University of Louvain (ref. no. B403201111965).

Stimuli

Natural face images of 12 male Caucasian celebrities taken from the internet were used as stimuli. Six of the celebrities were used as familiar faces (*Dany Boon, Emmanuel Macron, Garou, Jean Dujardin, Nagui,* & *Nicolas Sarkozy*), and another 6 as unfamiliar faces (*Andrzej Piaseczny, Cristi Puiu, Danny Dyer, Dermot Oleary, Pawel Delag,* & *Robert Biedron*). We used pictures of celebrities in both cases because they appear in similar contexts and poses, but our Belgian French-speaking participants were familiar only with the faces of the first pool of celebrities (which, but for one identity, was the same as used in the validation of the paradigm with clear images; Yan & Rossion 2020). This was verified by a Face Questionnaire acquired after EEG testing (Yan et al. 2020; Yan & Rossion 2020; Zimmermann et al. 2019). All participants were very familiar with the faces of the celebrities in the first pool. Only one participant could not recognize the face of *Jean Dujardin*, and another one had trouble recognizing *Nagui*. Two other participants could recognize the face of *Nagui*, but not his name. None of the participants could recognize the unfamiliar celebrity faces. We selected 20 different natural face images of each identity, yielding a total of 120 familiar and 120 unfamiliar face images. Faces were embedded in their natural background, varying greatly in orientation, lighting, and overall appearance (even among different images of the same identity) (Fig. 1; see also Yan & Rossion 2020). All face stimuli were cropped into a squared shape with 200 pixels in height/width and transformed into grayscale. When viewed from 50 cm away, each image extended to a visual angle of approximately 9.72°. The mean rectangular region across all 240 face images delimited by the head (including hair and ears) subtended about a visual angle of 8 (± 0.65)° × 6.2 (± 0.7)° (approximately covering an area of 52.6% ± 9.8% of the image surface).

We applied a low-pass spatial filter on the images at 9 logarithmic cut-off steps (referred to here as SF steps; Quek et al. 2018). Since most of the image energy is contained in LSF ranges for natural images, in particular for faces compared with other nonface categories (e.g., vehicles, animals; Guyader et al. 2017; Keil 2008), we used logarithmic instead of linear cutoffs. For image processing, we first normalized the full-spectrum grayscale images to obtain a global luminance with 0 mean and a standard deviation (SD) equal to 1. Subsequently, filtered images were obtained by fast Fourier transforming each image and multiplying the Fourier energy with the Gaussian filters. The images were low-pass filtered from 3 to 40 cycles/image (cpi) with 9 steps (Fig. 1). This resulted in a total of 2160 images across SF steps. Following spatial filtering, the luminance and contrast of all face images from each SF step were adjusted to match the mean values of the original full-spectrum image set to guarantee equal global luminance and contrast value both within and across SF steps (Quek et al. 2018). Figure 1 also indicates the SF cutoffs in terms of cycles/degree. Most importantly, here, since we used natural heterogenous face images without cropping

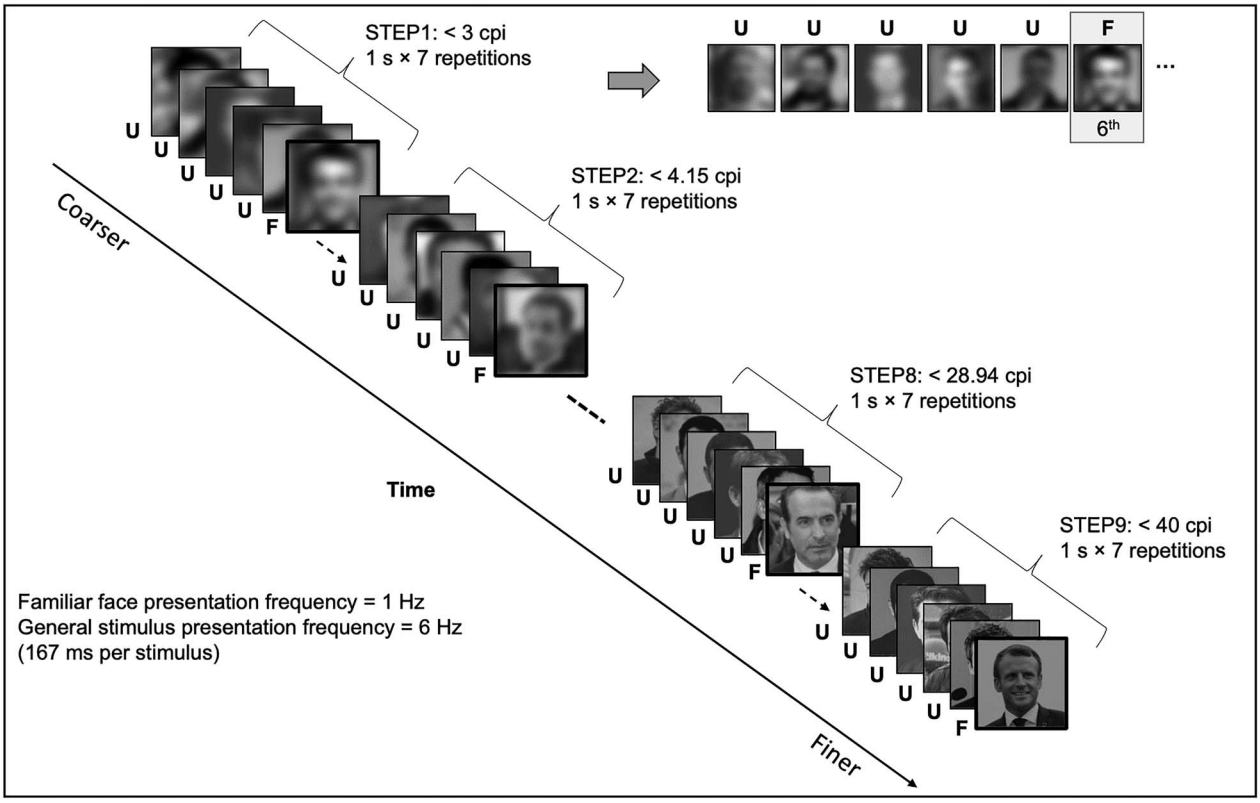

Figure 2. Schematic illustration of the coarse-to-fine (CF) presentation condition with parametrically increasing SF content at every 7 s throughout the sequence over 63 s (with examples showing the first 2 and last 2 SF steps) (see Supplementary Fig. S1 in supporting information for the fine-to-coarse condition). Unfamiliar (U) faces were presented at a fixed rate of 6 images by second (i.e., 6 Hz) with different familiar (F) face identities embedded at every 6th image (i.e., 1 Hz). During each sequence, images were presented through square-wave contrast modulation with no ISI. Face images shown here in the last 2 SF steps as familiar faces are with license permits. For license information, Jean Dujardin: Pictures licensed under the Creative Commons Attribution-Shared Alike 3.0 Unported. Attribution: Georges Biard. Emmanuel Macron: public domain.

out external features such as hair and ears, which contribute significantly to FFR, we report SF values in cycles/head width, based on the mean head width across all familiar face images. Since head size varied substantially across face images, we also report the SD of cycles/head width (Fig. 1A).

Procedure

EEG Testing

The same frequency-tagging method of Yan and Rossion (2020) was used to measure FFR. Images of unfamiliar faces were presented at a fixed rate of 6 Hz (i.e., 6 images by second) during each 63-second sequence, with different familiar faces inserted at every sixth image (i.e., 1 Hz). Then we used a sweep design to systematically modulate the image SF contents during each sequence (Ales et al. 2012; Quek et al. 2018). To be more specific, in 1 condition, the SF content of the images gradually increased every 7 s over the course of 9 sequential SF steps (Fig. 2). In this way, the images initially appeared blurry (i.e., low-pass filtered), gradually sharpening into full-spectrum image over the course of 63 s (referred to here as coarse-to-fine, CF condition). In the other presentation condition, inversely, the SF content of images gradually "decreased" over the course of 9 sequential SF steps (see Fig. S1). Therefore, the images in the sequence initially appeared clearly, with very fine details, gradually becoming blurry over 63 s (referred to here as fine-to-coarse, FC condition). Sequences of these 2 presentation conditions were flanked by 7-second pre- and post-ludes which were repetitions of the first and last SF steps, respectively, in order to reduce the eye-movements and muscular artifacts induced by the abrupt onset and offset of flickering stimuli. The total stimulation duration of a full sequence was 77 s. During each stimulation cycle, the face images were presented through square-wave contrast modulation with a 100% contrast of each cycle. In total, participants had to complete 20 CF and 20 FC sequences with a random order across 2 blocks (each containing 10 CF and 10 FC sequences).

During each stimulation sequence, the face images were presented at the center of the screen, with a vertical bar on each side of the image. The size of the vertical bar was 200 pixels (height) × 4 pixels (width), with a visual angle subtending 9.72 × 1.72° at a distance of 50 cm. The vertical bars appeared at the beginning of each sequence and disappeared until the sequence ended. The color of each bar changed 10 times during each sequence (from blue to pink, 2-second duration). Participants were required to press a space bar as soon and as accurately as possible when they saw the 2 bars changing color simultaneously (5 times by average), while at the same time being instructed to monitor the face images on the screen. The whole experiment took approximately 90 min, including breaks.

EEG Acquisition

The experiment was run in a quiet and low-lit room. The stimulation sequences were presented on a LED monitor (BenQ XL2420T) with a 1600×900 window resolution and a 120 Hz refresh rate. Stimuli were presented centrally on the screen. High-density 128-channel EEG was acquired with the ActiveTwo Biosemi system (Biosemi, Amsterdam, the Netherlands) at a 512 Hz sampling rate. The magnitude of the offset of all electrodes, referenced to the common mode sense, was held below 30 μV. Vertical and horizontal electrooculogram was recorded using 4 additional flat-type active-electrodes: 2 electrodes above and below the participant's right orbit and 2 lateral to the external canthi of the 2 eyes.

Analysis

EEG Preprocessing

EEG data analysis was carried out in the free software Letswave 5 (http://www.nocions.org/letswave) running on Matlab R2018a (MathWorks, USA). EEG data were first band-pass filtered between 0.05 and 100 Hz with a fourth order zero-phase Butterworth filter and then down-sampled to 256 Hz to reduce computational load. The data sequence was then segmented relative to the starting trigger of each trial, with an additional 2 second before and after each sequence (−2 to 79 s). To correct for eye blinks, an independent component analysis was applied for 3 participants who blinked more than 0.2 times/s on average throughout the testing sequences (0.15 ± 0.13 blinks/s; Retter & Rossion 2016). Individual channels with artifacts were interpolated by their 3 neighboring channels. The maximum number of interpolated channels for each participant was 6 (1.6 ± 1.8 channels). The cleaned-up data was then referenced to the average of all 128 channels.

EEG Frequency Domain Analysis

The preprocessed data were cropped again into epochs according to each SF step (9×7 s epochs). The first (prelude) and last (postlude) 7 s of each stimulation sequence were discarded to remove eye-movements and transients related to the abrupt onset and offset of the flickering stimuli. A Fast Fourier Transform was applied to each averaged SF step epoch and the amplitude spectra were extracted, with a frequency resolution of 0.143 Hz (1/7 s). To correct for variations in baseline noise level around each frequency of interest, 2 methods were used: (1) the mean amplitude of the neighboring 6 bins (3 by each side) were calculated and divided from each target frequency bin to display EEG spectra in signal-to-noise ratio, allowing better visualization of small responses (Retter & Rossion 2016) and (2) subtraction of the EEG noise from target bins (baseline subtraction) to quantify responses in microvolt. At each SF step, the FFR response was quantified with a summation of response at 1 Hz and harmonics up to 9 Hz (excluding 6 Hz that coincided with the general visual stimuli presentation frequency) (Retter & Rossion 2016; see e.g., Yan, et al. 2020; Zimmermann, et al. 2019). The general visual response was quantified with a summation of the first 8 harmonics at 6 Hz (i.e., up to 48 Hz).

To investigate FFR as a function of SF steps, a 10-channel bilateral occipitotemporal (OT) region of interest (ROI) was defined as a priori, in line with previous studies that showed the maximum responses to faces versus nonface objects (Retter & Rossion 2016; Quek et al. 2018) or to familiar compared with unknown faces (Yan, et al. 2020; Yan & Rossion 2020; Zimmermann, et al. 2019). The OT ROI included channels P7, P9, PO7, PO9, PO11 over the left hemisphere, and channels P8, P10, PO8, PO10, PO12 over the right hemisphere. The general visual response was observed over a comparatively larger ROI centered over the most posterior electrodes. Based on the general visual response topographies across SF steps and conditions, we defined a large posterior ROI including 16 channels in the bilateral OT region (i.e., P5&6, P7&8, P9&10, PPO3&4, PPO5&6, PO7&8, PO9&10, and PO11&12) and 15 channels (i.e., PO3&4 h, POO5&6, O1&2, POI1&2, I1&2, POz, POOz, Oz, OIz, and Iz) in the medial occipital region.

Further data analysis mainly includes 3 parts. First, parametric statistical analyses were carried out using repeated measured ANOVAs for both FFR and the general visual responses. Greenhouse–Geisser corrections were applied to degrees of freedom whenever the assumption of sphericity was violated. Pairwise *t*-tests were used for post-hoc comparisons and Bonferroni corrections were applied for multiple comparisons. Second, we measured the FFR threshold (i.e., significantly above 0) in each presentation condition (i.e., CF and FC) with a bootstrap method: (1) 16 data points (i.e., 16 participants) were randomly drawn with replacement from each SF step and the mean of the drawn points was computed; (2) the procedure was repeated 5000 times; (3) the mean and 95% confidence interval of the resulting distribution from step (2) were calculated. The recognition threshold was defined at the first SF step (i.e., in the CF condition) when the confidence interval did not overlap with 0. Third, we modeled the relationship between the response amplitudes and SF steps to describe the response patterns of the general visual responses of each presentation condition. We built a power function ($y = a \times x^b$) and a linear function ($y = a \times x + b$) separately for both conditions. In both cases, y indicates the response amplitudes and *x* indicates the SF steps.

Results

The Neural FFR Response as a Function of SF Content

Figure 3 shows the averaged frequency spectra of FFR at 1 Hz (and harmonics) across the 2 presentation conditions at each SF step over the OT ROI. Significant harmonics (above EEG noise) emerged at step 5, when images were low-pass filtered at 10.95 cycles/image (cpi, corresponding to about 7.01 ± 0.82 cphw, i.e., <8.65 cphw for 96% of the images). The recognition response mainly locates over the posterior region of the head, especially over bilateral OT region.

To investigate the FFR as a function of SF steps at different presentation conditions, we first ran a repeated-ANOVA with Condition (CF, FC) and SF step as within-subjects factors across all 128 channels averaged together. The results only showed a significant main effect of *SF step*, $F(8,120) = 3.24$, $P < 0.01$, $\eta^2 = 0.18$. No other effects were significant (both $Ps > 0.1$). We then ran a repeated-ANOVA with Hemisphere (Left, Right), SF step, and Condition (CF, FC) as within-subjects factors over the OT ROI. Again, we only found a significant main effect of *SF step*, $F(8,120) = 10.52$, $P < 0.001$, $\eta^2 = 0.41$. The main effect of Hemisphere was not significant, $F(1,15) = 2.6$, $P > 0.1$, $\eta^2 = 0.02$, indicating that there was no hemispheric difference over the OT ROI (Fig. 4). No other effects were significant (all $Ps > 0.1$).

Although the ANOVA did not show effects of presentation conditions on response amplitudes, clear differences in response patterns across SF steps were visible between the 2

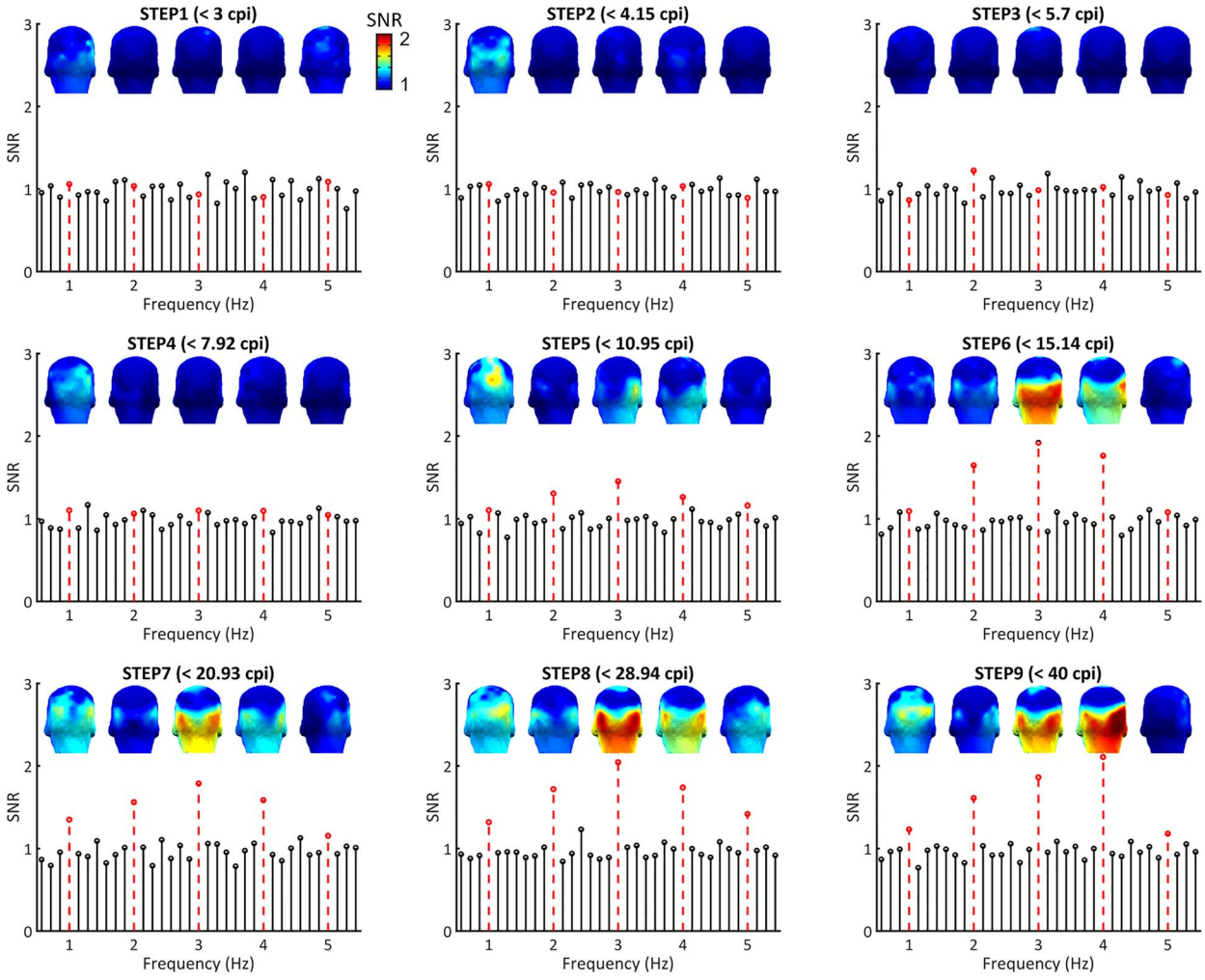

Figure 3. Grand-averaged frequency spectra in signal-to-noise ratio across conditions as a function of SF steps over the OT ROI for FFR. The frequency bins at every 1 Hz step (up to 5 Hz) are highlighted in red dashed lines. Three-D topographies of the first 5 harmonics of 1 Hz are shown above their corresponding harmonics.

presentation conditions (Fig. 4). To investigate these observations, we used a bootstrap statistical method to determine the recognition threshold (i.e., starting to be significantly above 0) at each presentation condition (see Methods part). For the CF condition, familiar faces could be reliably recognized already at step 5 (i.e., 10.95 cpi). When considering SF information according to head width, this step 5 corresponds to a low-pass filter below 8.65 cycles/head width (cphw) for most of the familiar faces (115 of 120, i.e., 96%; Fig. 5A). In contrast, for the FC condition, the recognition threshold shifted by one step (step 6), that is, when images were low-pass filtered at 15.14 cpi (i.e., <11.97 cphw for 96% of the images). In addition, the recognition response to faces low-pass filtered at 10.95 cpi (i.e., <8.65 cphw) was significantly larger in the CF condition than that in the FC condition ($t_{(15)} = 2.94$, $P < 0.01$). Familiar and unfamiliar face exemplars that were filtered one step below or at 10.95 cpi are shown in Figure 5B.

In both the CF and FC conditions, the increase of EEG amplitude (as the SF content accumulated), marking the onset (in terms of SF cutoff) of the FFR neural response was abrupt, with the remaining increase in SF content having little or no effect on the response amplitude (Fig. 4). Supporting this observation, an ANOVA restricted to SF steps 5–9 in the CF condition failed to show any effect of frequency steps ($P > 0.1$), and an ANOVA restricted to SF steps 6–9 in the FC condition did not show any effect of frequency steps, either ($P > 0.1$).

General Visual Response as a Function of SF Steps

The general stimulation rate of 6 Hz elicited significant general visual responses at the first 8 consecutive harmonics (i.e., 6 Hz, 12 Hz, and so forth) across all participants and presentation conditions for each SF step, with this response located mainly over the posterior regions. In general, the overall response amplitude increased gradually as the SF content of the images increased from 3 to 40 cpi (Figure S3B). The spatial distribution of neural responses shifted from the middle occipital regions to a larger posterior area including both the middle occipital and bilateral OT regions when more SF information was presented.

There was a difference in response patterns between the 2 temporal presentation conditions: in the CF condition, the response increased only slightly (34%) and smoothly from the coarsest stimuli to the highest resolution stimuli. In contrast, in the FC condition, the response dropped much

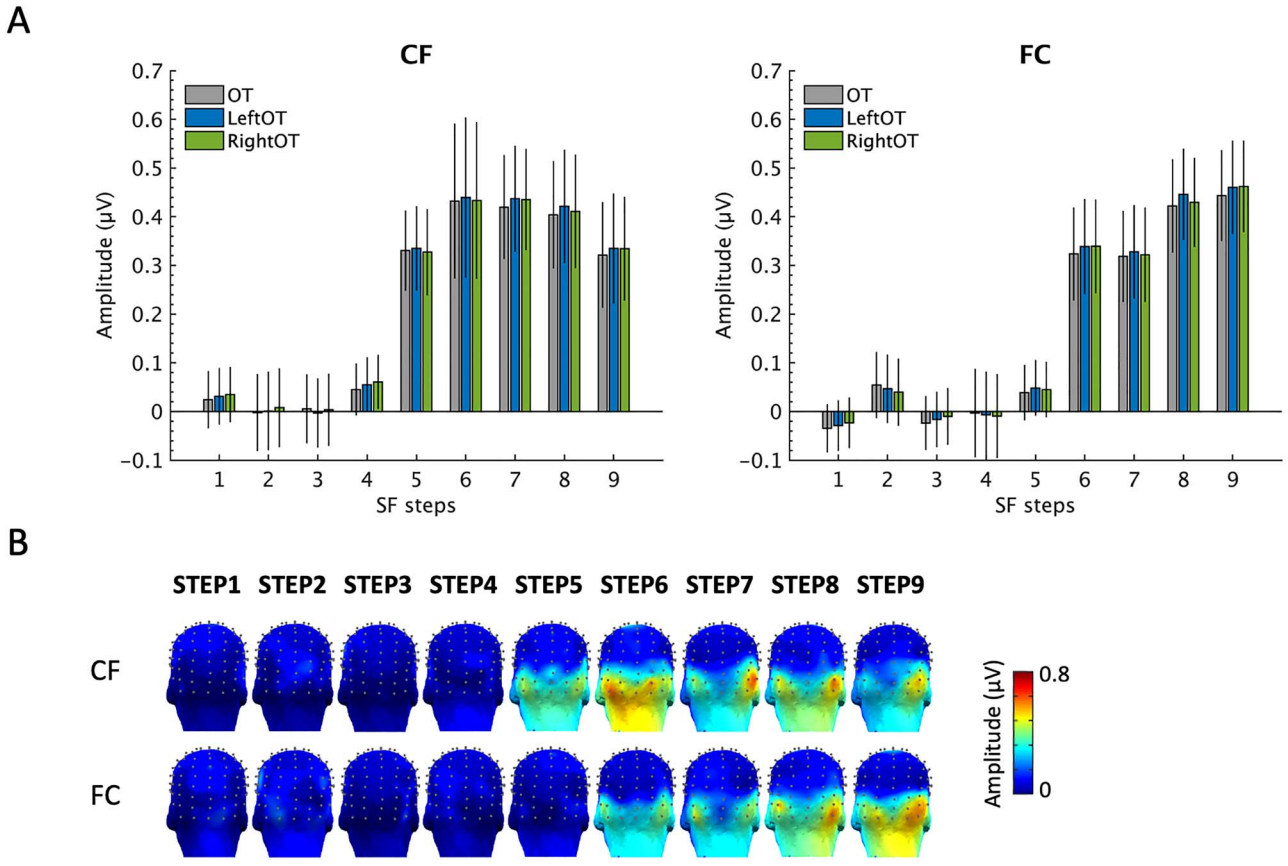

Figure 4. FFR response. (A) Grand-averaged FFR as a function of SF steps (i.e., from 3 to 40 cpi) at the 2 presentation conditions (CF, left panel; FC, right panel) over different ROIs (left OT, right OT, and OT). Error bars indicate standard error of the mean. (B) Three-D scalp topographies averaged across all participants of each SF step.

more largely (85%) and linearly as finer cues disappeared from the stimuli (Fig. 6). These patterns were further modeled with the response amplitudes and the SF steps for each presentation condition over the posterior ROI. The response pattern of the CF condition fitted well with a power function, $y = 3.4 \times x^{0.1}$-2 (adjusted $R^2 = 0.939$), better than a linear function, $y = 0.07 \times x + 1.5$ (adjusted $R^2 = 0.827$). In the FC condition, the response pattern fitted comparably well to a linear function, $y = 0.14 \times x + 1.2$ (adjusted $R^2 = 0.995$), or to a power function, $y = 0.16 \times x^{0.1} + 1.2$ (adjusted $R^2 = 0.993$).

We also ran repeated-ANOVAs with SF step and Condition as within-subjects factors on the summed-harmonic baseline-corrected response amplitudes across all 128 channels and over the posterior ROI, separately. There was no effect of Condition (both $Ps > 0.1$). However, there were significant main effects of SF step (all channels: $F(8,120) = 29.38$, $P < 0.001$, $\eta^2 = 0.66$; posterior ROI: $F(8,120) = 26.88$, $P < 0.001$, $\eta^2 = 0.64$). The interaction of SF step × Condition was also significant (all channels: $F(8,120) = 4.9$, $P < 0.001$, $\eta^2 = 0.25$; posterior ROI: $F(8,120) = 3.73$, $P < 0.001$, $\eta^2 = 0.2$). Further analyses indicated that for both presentation conditions, there were significant main effects of SF step (both $Ps < 0.01$).

Discussion

We took advantage of a recently developed frequency-tagging paradigm to measure automatic FFR under time constraints (i.e., single glance) in order to define the level of spatial resolution necessary and sufficient to perform this key function for the human brain. As in previous studies with unfiltered images, the neural FFR response was objectively identified in the EEG spectrum at 1 Hz (and harmonics), over bilateral OT channels (Yan & Rossion 2020). The lack of significant right hemispheric lateralization—typical of face recognition in the human species—has been discussed in previous studies and attributed to the definition of the response of interest here, which reflects a direct contrast between familiar and unfamiliar faces (Campbell et al. 2020; Yan et al. 2020; Zimmermann et al. 2019).

Coarse Information (below 8.65 Luminance Cycles/Head Width) Suffices for Single-Glanced Automatic FFR

In the coarse-to-fine stimulation mode, the neural index of FFR emerged significantly over OT regions for images that were low-pass filtered with a 10.95 cpi. For our image set with natural variable images, this value corresponds to 7.01 ± 0.82 cycles by head width. Below this range (i.e., step 4, 5.07 ± 0.6 cphw), no significant FFR response was recorded at all (Fig. 3). Hence, our study shows that the automatic, single-glanced, recognition of a natural image of a face as being familiar (here famous) among unknown faces requires more than 6.27 (i.e., 5.07 + 2SD) and up to about 8.65 (i.e., 7.01 + 2SD) variations of luminance by head width.

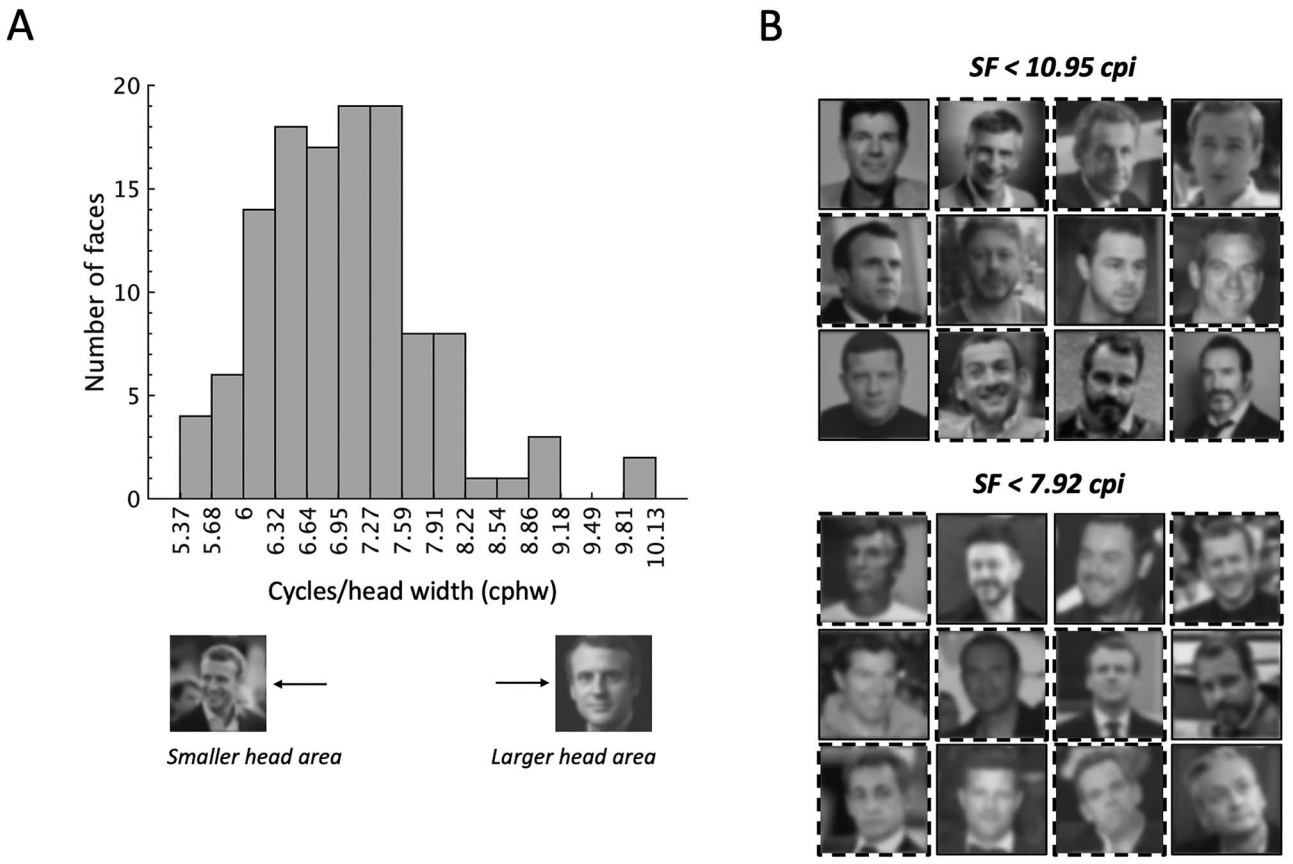

Figure 5. Stimuli exemplars. (*A*) Histograms showing the SF information distribution of all 120 familiar face images in cycles/head width (cphw) converted from the 10.95 cycles/image (cpi). For 96% (i.e., 115 of 120) of the stimuli low-pass filtered at 10.95 cpi, the SF cutoff corresponds to less than 8.65 cphw. (*B*) A few face exemplars (familiar faces highlighted with dashed outlines) low-pass filtered at 10.95 cpi (i.e., step 5, top panel, a level at which a full FFR response emerged in the CF condition), and at 7.92 cpi (bottom panel). See Supplementary Figure S2 for the same face exemplars low-pass filtered at the reversed cutoffs (i.e., faces in the left panel were low-pass filtered at 7.92 cpi, and in the right panel at 10.95 cpi).

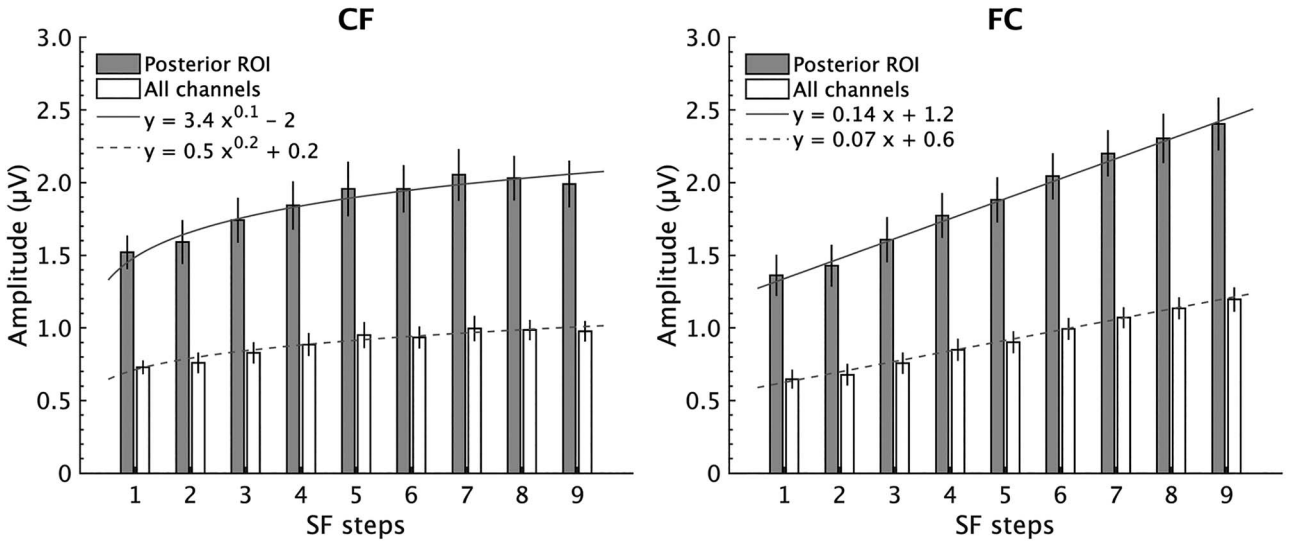

Figure 6. General visual responses. Grand-averaged general visual responses as a function of SF steps at 2 presentation conditions (CF, left panel; FC, right panel) across all channels and over the posterior ROI, with model estimation of the relationship between the response amplitudes and SF steps. Error bars indicate standard error of the mean. Solid and dashed lines indicate corresponding estimated function with the model.

This result was obtained with head widths of about 6.2° on average in our stimulus set. For a (Caucasian male) head width of about 15 cm (Farkas 1994), this corresponds to a real face seen at about 1.39 m of distance, as experienced most frequently in real life circumstances (Oruc et al., 2019). In terms of SF cycles, our identified threshold therefore corresponds to at least 1.4 cycles/degree of visual angle. However, (familiar) face identity recognition is invariant across a wide variety of stimulus size (Andrews & Ewbank 2004; Lee, Matsumiya, & Wilson 2006; Oruç& Barton 2010; Zhao & Chubb 2001), so that—unless the face becomes too small to be resolvable by human visual acuity (see below)—the relevant value here is the threshold as expressed in number of cycles/face (Hayes, Morrone, & Burr 1986; Ginsburg 1980; Loftus & Harley 2005; Millward & O'Toole 1986; Parker & Costen 1999; see also Morrison & Schyns 2001) or cycles/head width as used in the present study. Although increasing stimulus size (i.e., decreasing distance) should not lower the threshold identified for (upright) faces (Oruç& Barton 2010), we cannot exclude that decreasing stimulus size may slightly improve the response at 8.65 cycles/head width (Ojänpää & Nasanen 2003; Oruç& Barton 2010; Shahangian & Oruç2014; Mousavi & Oruc 2020; but see Hayes et al. 1986) or even lead to a significant response for the lower step value (5.07 ± 0.6 cphw, Fig. 5A) used here. However, the lack of any significant response at this lower cutoff in our EEG study makes it unlikely. Moreover, if the stimulus becomes too small (i.e., too far away), the value of 8.65 cycles/head width may fall below the limit of human visual acuity. In fact, other factors than stimulus size could certainly play a more significant role to lower this threshold. For instance, the threshold could potentially be lowered if pictures of the familiar faces were less variable, in full color, shown to the observers before running the study, more frequently repeated, and not directly contrasted with variable pictures of unfamiliar faces acting as forward- and backward-masks in the stimulation sequence (Fig. 1).

Using the same sweep frequency-tagging approach, Quek et al. (2018) reported that human observers need approximately 1.66 cycles/face width to recognize a visual stimulus as a face among other nonface objects. Therefore, our findings indicate that a substantially higher level of spatial resolution is required for face familiarity recognition. This difference is not surprising given the much higher physical similarity between the distractors (i.e., either various object shapes or unfamiliar faces) and the target stimuli (i.e., either faces or familiar faces) in the respective measures. The present observation is also in line with a wealth of evidence indicating that accumulation of sensory signals in the human brain reaches a threshold earlier for recognizing a face as a face than for recognizing it as being familiar or accessing its identity (Sergent 1986; see Amihai et al. 2011; Barragan-Jason, Lachat, & Barbeau 2012; Besson et al. 2017 for direct comparisons).

The neural threshold of 8.65 cycles/head width as identified here for the first time with physically variable stimuli roughly agrees with a number of behavioral studies that reported the largest drops of face recognition performance below about 8 cycles/face (Bachmann 1991; Bachmann & Kahusk 1997; Costen et al. 1994, 1996; Fiorentini et al. 1983; Gold et al. 1999; Hayes, Morrone, & Burr 1986; Oruç& Barton 2010; Parker & Costen 1999; Peli et al. 1991; Peli et al. 1994). Although there are early reports of famous faces with external features being well recognized even at lower spatial resolution (Harmon 1973; Harmon & Julesz 1973; Ginsburg 1980; Rubin & Siegel 1984; see also Sinha 2002; Yip & Sinha 2002), there was usually no systematic data recorded in these reports and, most importantly, recognition was limited to a few iconic pictures of faces (e.g., the quantized picture of Lincoln in Harmon 1973; Harmon & Julesz 1973; see also the face celebrities in Fig. 1 of Sinha 2002).

FFR Saturates at 8.65 Cycles/Head Width

Another important, and striking, aspect of the present findings is that, in the coarse-to-fine presentation mode, the neural FFR response already saturated at this low-pass cutoff (i.e., <8.65 cycles/head width), that is, it did not increase further when images increased in spatial resolution (Fig. 4A, left panel). Similarly, in the fine-to-coarse stimulation mode, the drop of the FFR response was equally abrupt (Fig. 4A, right panel), albeit at a higher cut-off value. This observation is in line with studies as mentioned above and in the introduction (Bachmann 1991; Costen et al. 1994; Oruç& Barton 2010; Peli et al. 1994) and is difficult to reconcile with the prevalent view in the scientific literature that a medium range of SF, of about 8–16 cycles/face, would be optimal for face identity recognition (i.e., "The golden mean"; e.g., Collin et al. 2006; Costen et al. 1994, 1996; Fiorentini et al. 1983; Gao & Maurer 2011; Gold et al. 1999; Keil 2008; Näsänen 1999; Parker & Costen 1999; Ojänpää & Näsänen 2003; Schyns et al. 2002; see also Collin et al. 2012 for human electrophysiological evidence). A number of reasons could be advanced to explain the significantly lower threshold value found here.

First, to elicit a significant neural FFR response, the present paradigm does not require access to the specific identity of the familiar face, but only to distinguish a familiar face (irrespective of its specific identity) from unfamiliar faces. Yet, to access familiarity, each familiar face must be distinguished at the individual level from other (unfamiliar) faces presented in the sequence so that this difference is unlikely to play a significant role in the low threshold observed here.

Second, and most importantly, studies that support the key contribution of the medium SF range of about 8–16 cycles/face usually use homogenous face stimuli in which the external features, such as the hair (and sometimes the ears), are removed (Bachmann 1991; Gao & Maurer 2011; Gold et al. 1999; Hsiao et al. 2005; Näsänen 1999; Tanskanen et al. 2005; Ramon et al. 2015; see also Morrisson & Schyns 2001 for review). From a methodological point of view, the removal of external features and facial hair in stimulus sets is valid and important because these studies usually rely on unfamiliar face matching or old/new face recognition tasks, with only one full-front image per face identity. However, facial hair and external features constitute relatively coarse diagnostic cues of identity (Abudarham & Yovel 2016; Sinha & Poggio 1996) and their removal undoubtedly increases stimulus homogeneity and impairs identity recognizability (Bonner, Burton, & Bruce 2003; Johnston & Edmonds 2009). Hence, these manipulations are likely to promote the reliance on diagnostic cues available at higher SF ranges than the 8.65 cycles/face width threshold identified in the present study. Here, importantly, since our neural measure relies on the recognition of previously unseen and highly variable numerous (20) images of different familiar identities against variable distractors, the presence of coarse external features or facial hair is not a confounding factor. In addition, the face stimuli used in the present study are not segmented from their natural background and vary substantially in viewing conditions (e.g., head orientation, size, lighting), making their recognition particularly challenging.

Third, contrary to most of the previous studies (but see Bachmann 1991; Bachmann & Kahusk 1997; Ramon et al. 2015), the SF threshold for FFR is estimated here in a "dynamic" SF display rather than with static images presented at different cutoffs. A major finding of the present study, indeed, is that the temporal dynamics matters, since the 8.65 cycles/head width threshold was reached only in the coarse-to-fine presentation mode. Had we used only a fine-to-coarse stimulation mode, our conclusions regarding the SF threshold would have been relatively more in line with the outcome of behavioral studies. This point is further discussed below.

Fourth, our measure is taken in challenging conditions, allowing only 166 ms of display per face (i.e., one gaze fixation) and with each face being forward- and backward-masked by other faces. Although this is in line with evidence for rapid, single-glanced, recognition of facial identities (Barragan-Jason et al. 2015; Caharel et al. 2014; Hacker et al. 2019; Hsiao & Cottrell 2008; di Oleggio Castello & Gobbini 2015; Yan & Rossion 2020), and can be compared with the ecological experience of automatically and unexpectedly recognizing a familiar face in a dynamic crowd of strangers, this stimulation mode certainly prevents a slower detailed analysis of facial features relying on higher SF ranges.

Finally, an obvious difference with previous studies is that an implicit, or task-free, neural measure is recorded here, as opposed to explicit behavioral measures in previous studies. As mentioned above, in the previous study providing a neural (fMRI) measure of human FFR in a coarse-to-fine display (Ramon et al. 2015), an explicit behavioral task was used. Although an average RT of 10.09 cycles/face was found for familiar decisions, faces were classified as unfamiliar only after 16.97 cycles/face, showing again that decisional factors play a significant role in the estimation of thresholds in such explicit behavioral tasks (see also Bruner & Potter 1964).

In summary, our task-free EEG measure with natural unsegmented images suggests that in the presence of SF up to 8.65 cycles/head width, additional "medium-range" SF are redundant for FFR, supporting the conclusions of early behavioral studies (Bachmann 1991; Costen et al. 1994; Peli et al. 1994; Oruç & Barton 2010) and extending them to the automatic single-glanced, recognition of natural (unsegmented) heterogenous views of familiar faces.

Overall, these findings agree with clinical observations according to which low-level visual defects, in particular low visual acuity, does not prevent familiar face (identity) recognition in the real world and preserves this function compared with the spectacular impairment observed in brain-damaged cases of prosopagnosia (de Haan et al. 1995). In the same vein, although a decrease of contrast sensitivity of high spatial frequency in the normal elderly population may cause poor face identity recognition performance in certain tasks (e.g., Cronin-Golomb et al. 2007; Owsley, Sekuler, & Boldt 1981), the ability to recognize familiar faces in this population should be preserved if evaluated with the present paradigm. More generally, our observations provide useful and objective information to help assessing a high-level ecological visual recognition function such as FFR, in populations with low-level visual impairments at adulthood and also during development, normal or pathological aging (e.g., cortical visual impairment; amblyopia; glaucoma; age-related macular degeneration; Glen et al. 2012; Lane et al., 2018a, 2018b).

Moreover, since relatively low resolution images are sufficient for FFR even in challenging stimulation conditions, this finding has important practical implications for developing valid face recognition and image compression algorithms (Adjabi et al. 2020; Bachmann 2016; Goffaux 2016; Strang & Nguyen 1996).

At the theoretical level, these observations suggest that visual representations of familiar faces are quite coarse (Sergent 1986; Sinha 2002). Such coarse visual representations may provide substantial advantages in terms of neural coding, making it resistant to noise and distortions (Hole et al. 2002), which is key to efficient FFR (Burton 2013; Young & Burton 2017). Matching low-level sensory inputs to coarse visual representations of faces is also much more efficient if these representations are holistic, that is, undecomposed in parts or features (Collishaw & Hole 2000; Goffaux & Rossion 2006; Rossion 2013; Sergent 1986; Tanaka & Farah 1993; Young et al. 1987). Indeed, although isolated facial parts, or parts that are presented in disrupted configurations, of previously seen images can be recognized above chance level when presented at high resolution, recognition of such isolated/disconfigured parts would be extremely challenging if not impossible at a coarse level of resolution as identified here (Schwaninger et al. 2002).

An All-or-None FFR Response?

A logical conclusion of the points discussed in the 2 sections above is that the threshold and saturation points for FFR are the same, suggesting an all-or-none recognition response. This abrupt neural response function differs from the relatively more gradual increase in amplitude observed in the general visual response common to all faces (i.e., familiar and unfamiliar faces) over posterior electrodes (Fig. 6). Most importantly, it differs from the face-selective response obtained (over the same OT electrodes) from human observers when natural images of faces are progressively revealed in a coarse-to-fine display among images of nonface objects (with similarly space SF steps; Quek et al. 2018). Although this gradual increase was interpreted in terms of progressive accumulation of evidence for recognition, it may also reflect an abrupt, all-or-none visual recognition of each individual stimulus in the sequence, the overall gradual increase in response amplitude reflecting the increasing number of stimuli being recognized (in an all-or-none fashion) in the sequence (Retter et al. 2020). Regardless of the correct account for the previous observations in generic face recognition, the all-or-none FFR neural response identified here suggests that the progressive accumulation of sensory information takes place in low-level visual regions before matching representations of familiar faces in higher order regions of the visual cortex. Once this match is registered, a threshold for recognition is reached, and unless specific explicit analysis of the face identity would be required, further sensory information at higher SF does not need to be added.

The Coarse-to-Fine Advantage for FFR

A second major finding of the present study is that recognition of familiar faces was affected by the temporal dynamics of presentation mode, since different recognition thresholds were reached in different conditions. That is, compared with the fine-to-coarse condition, fewer spatial frequency details (i.e., 8.65 vs. 11.97 cycles/head width) were required to elicit a full recognition response in the coarse-to-fine condition. Interestingly, this type of asymmetry between coarse-to-fine versus fine-to-coarse SF

stimulation conditions in electrophysiological recordings has not been observed with low-level visual stimuli (e.g., gratings or checkerboards to measure visual acuity; Almoqbel et al. 2011; Hemptinne et al. 2018; see Hamilton et al. 2021 for review).

The coarse-to-fine advantage in familiarity recognition also seems at odd with the perceptual hysteresis effects obtained with SF-filtered faces (Brady & Oliva 2012; see also Bruner & Potter 1964). In Brady & Oliva (2012), participants were asked to recognize approaching (i.e., coarse-to-fine) versus receding (i.e., fine-to-coarse) hybrid displays (i.e., superimpositions of different LSF and HSF images). It is only in the receding, fine-to-coarse, condition, that perceptual hysteresis took place with participants being able to hold previous experience of stimulus details to make inferences about the receding (progressively coarser) input stimulus. In contrast, in the approaching coarse-to-fine condition, there was no advantage of previously seeing a coarser version of the stimulus.

Two reasons could explain the discrepancy between these observations and the coarse-to-fine advantage of the FFR as observed here. First, we tested FFR across different identities, and with different, widely variable, natural face images for each identity. This heterogeneity in the stimulus set prevented participants' reliance on nonidentity-related image-based cues, whereas the use of the exact same face images in hysteresis and hindsight studies likely encouraged image-based strategies. Second, since we used different familiar face identities in each sequence with a random presentation at every fixed sixth image, together with a rapid and challenging presentation mode, the visual system might be hindered to make a rapid hypothesis based on previous experiences. Regardless, our findings do not support a hypothesis of the predictive coding framework according to which hypotheses based on previous experiences are continuously updated to predict the upcoming percept (Friston 2005; Rao & Ballard 1999; see also Trapp et al. 2021). Indeed, according to this framework, since the identity information gained with fine-detailed faces at the beginning of each sequence could be used to predict the upcoming more blurred images, a lower recognition threshold could have been anticipated in the fine-to-coarse condition. Conversely, our results are in line with the prevalence of coarse-to-fine sequencing of sensory information processing in the human brain for complex stimuli such as faces (Bachmann 1991; Hegdé 2008; Goffaux et al. 2011; Petras et al. 2019; Parker et al. 1992, 1997; Sergent 1986; Vogelsang et al. 2018; Watt 1987). In natural conditions, we see faces from further away before they come closer for recognition, but not, or rarely, the opposite (i.e., people moving away turn their back to us). Therefore, we rarely experience a fine-to-coarse presentation mode, in which fine details of faces are removed progressively. Moreover, it also reflects natural experience during development, since typically developing newborns commence their visual experience with remarkably poor visual acuity (Atkinson & Braddick 2013; Dobson & Teller 1978) due to immature neonatal retina (Banks & Bennett 1988) as well as to immaturities in the visual cortex (Jacobs & Blakemore 1988). Over the initial months of development, these immaturities diminish steadily, leading to acuity improvement and the integration of information at higher SF (Atkinson & Braddick 2013; Daw 2014), which may be particularly useful for the development of their normal ability to recognize facial identity based on visual inputs only (Le Grand et al. 2003; Vogelsang et al. 2018).

The coarse-to-fine view is also supported here by findings for the general visual response common to faces: although there was a smooth amplitude increase of this posterior response in the coarse-to-fine condition as more SF content of faces was accumulated, the response amplitude seemed to decrease more abruptly in the fine-to-coarse condition when faces became blurry (Fig. 6). This observation may suggest that spatial information in a coarse-to-fine sequence is integrated more efficiently than that in the reverse sequence. Most importantly, the abrupt emergence of the neural FFR response at an earlier step in the coarse-to-fine than in the fine-to-coarse mode provides original and particularly strong neural evidence supporting the coarse-to-fine processing view, showing that the automatic, single-glanced, recognition of a natural view of a familiar face in a temporal crowd of unfamiliar faces benefit from an accumulation of sensory cues in a naturally experienced order.

Supplementary Material

Supplementary material can be found at *Cerebral Cortex* online.

Funding

Region Grand Est (University of Lorraine to X.Y.), as well as by the Research Foundation - Flanders and Fonds de la Recherche Scientifique under the Excellence of Science (EOS) programme (HUMVISCAT-30991544 to V.G.). V.G. is a research associate of the National Fund for Scientific Research (F.R.S.-FNRS).

Notes

Conflict of Interest: The authors have declared that no conflict of interests exist.

References

Abudarham N, Yovel G. 2016. Reverse engineering the face space: discovering the critical features for face identification. *J Vis*. 16:1–18.

Adjabi I, Ouahabi A, Benzaoui A, Taleb-Ahmed A. 2020. Past, present, and future of face recognition: a review. *Electronics*. 9:1188.

Ales JM, Farzin F, Rossion B, Norcia AM. 2012. An objective method for measuring face detection thresholds using the sweep steady-state visual evoked response. *J Vis*. 12:1–18.

Almoqbel FM, Yadav NK, Leat SJ, Head LM, Irving EL. 2011. Effects of sweep VEP parameters on visual acuity and contrast thresholds in children and adults. *Graefes Arch Clin Exp Ophthalmol*. 249:613–623.

Amihai I, Deouell LY, Bentin S. 2011. Neural adaptation is related to face repetition irrespective of identity: a reappraisal of the N170 effect. *Exp Brain Res*. 209:193–204.

Andrews TJ, Ewbank MP. 2004. Distinct representations for facial identity and changeable aspects of faces in the human temporal lobe. *Neuro Image*. 23:905–913.

Atkinson J, Braddick O. 2013. Visual development. In: Zelazo PD, editor. *The Oxford handbook of developmental psychology*. Oxford (UK): Oxford University Press, pp. 271–309.

Bachmann T. 1991. Identification of spatially quantised tachistoscopic images of faces: how many pixels does it take to carry identity? *Eur J Cognit Psychol*. 3:87–103.

Bachmann T. 2016. *Perception of pixelated images*. Cambridge: Academic Press.

Bachmann T, Kahusk N. 1997. The effects of coarseness of quantisation, exposure duration, and selective spatial attention

on the perception of spatially quantised ('blocked') visual images. *Perception*. 26:1181–1196.
Banks MS, Bennett PJ. 1988. Optical and photoreceptor immaturities limit the spatial and chromatic vision of human neonates. *J Opt Soc Am A*. 5:2059.
Barragan-Jason G, Cauchoix M, Barbeau EJ. 2015. The neural speed of familiar face recognition. *Neuropsychologia*. 75:390–401.
Barragan-Jason G, Lachat F, Barbeau EJ. 2012. How fast is famous face recognition? *Front Psychol*. 30(3):454.
Besson G, Barragan-Jason G, Thorpe SJ, Fabre-Thorpe M, Puma S, Ceccaldi M, Barbeau EJ. 2017. From face processing to face recognition: comparing three different processing levels. *Cognition*. 158:33–43.
Bindemann M, Attard J, Leach A, Johnston RA. 2013. The effect of image pixelation on unfamiliar-face matching. *Appl Cognit Psychol*. 27:707–717.
Bonner L, Burton AM, Bruce V. 2003. Getting to know you: how we learn new faces. *Visual Cognit*. 10:527–536.
Brady TF, Oliva A. 2012. Spatial frequency integration during active perception: perceptual hysteresis when an object recedes. *Front Psychol*. 3:1–8.
Bruce V, Young A. 1986. Understanding face recognition. *Br J Psychol*. 77:305–327.
Bruner JS, Potter MC. 1964. Interference in visual recognition. *Science*. 144:424–425.
Burton AM. 2013. Why has research in face recognition progressed so slowly? The importance of variability. *Q J Exp Psychol*. 66:1467–1485.
Busigny T, Rossion B. 2010. Acquired prosopagnosia abolishes the face inversion effect. *Cortex*. 46:965–981.
Caharel S, Ramon M, Rossion B. 2014. Face familiarity: decisions take 200 msec in the human brain: electrophysiological evidence from a go/no-go speeded task. *J Cogn Neurosci*. 26:81–95.
Calder AJ, Rhodes G, Johnson MH, Haxby JV. 2011. *The Oxford handbook of face perception*. Oxford (UK): Oxford University Press.
Campbell FW, Cooper GF, Enroth-Cugell C. 1969. The spatial selectivity of the visual cells of the cat. *J Physiol*. 203:223–235.
Campbell A, Louw R, Michniak E, Tanaka JW. 2020. Identity-specific neural responses to three categories of face familiarity (own, friend, stranger) using fast periodic visual stimulation. *Neuropsychologia*. 141:107415.
Collin CA, Rainville SJM, Watier NN, Boutet I. 2014. Configural and featural discriminations use the same spatial frequencies: a model observer vs. human observer analysis. *Perception*. 43:509–526.
Collin CA, Therrien M, Martin C, Rainville S. 2006. Spatial frequency thresholds for face recognition when comparison faces are filtered and unfiltered. *Percept Psychophys*. 68:879–889.
Collin CA, Therrien ME, Campbell KB, Hamm JP. 2012. Effects of bandpass spatial frequency filtering of face and object images on the amplitude of N170. *Perception*. 41:717–732.
Collishaw SM, Hole GJ. 2000. Featural and configurational processes in the recognition of faces of different familiarity. *Perception*. 29:893–909.
Costen NP, Parker DM, Craw I. 1994. Spatial content and spatial quantisation effects in face recognition. *Perception*. 23:129–146.
Costen NP, Parker DM, Craw I. 1996. Effects of high-pass and low-pass spatial filtering on face identification. *Percept Psychophys*. 58:602–612.
Cronin-Golomb A, Gilmore GC, Neargarder S, Morrison SR, Laudate TM. 2007. Enhanced stimulus strength improves visual cognition in aging and Alzheimer's disease. *Cortex*. 43:952–966.
Daw NW. 2014. *Visual development*. 3rd ed. US: Springer.
De Haan EHF, Heywood CA, Young AW, Edelstyn N, Newcombe F. 1995. Ettlinger revisited: the relation between agnosia and sensory impairment. *J Neurol Neurosurg Psychiatry*. 58:350–356.
De Valois RL, De Valois KK. 1980. Spatial vision. *Annu Rev Psychol*. 3:309–341.
Di Oleggio Castello MV, Gobbini MI. 2015. Familiar face detection in 180ms. *PLoS One*. 10:e0136548.
Dobson V, Teller DY. 1978. Visual acuity in human infants: a review and comparison of behavioral and electrophysiological studies. *Vision Res*. 18:1469–1483.
Ellis A. 1975. Recognizing faces. *Br J Psychol*. 66:409–426.
Farkas LG. 1994. *Anthropometry of the head and face*. New York: Raven Press.
Fiorentini A, Maffei L, Sandini G. 1983. The role of high spatial frequencies in face perception. *Perception*. 12:195–201.
Friston K. 2005. A theory of cortical response. *Philos Trans R Soc B*. 360:815–836.
Gao X, Maurer D. 2011. A comparison of spatial frequency tuning for the recognition of facial identity and facial expressions in adults and children. *Vision Res*. 51:508–519.
Ginsburg A. 1980. Specifying relevant spatial information for image evaluation and display design: an exploration of how we see certain objects. *Proc SID*. 21:219–227.
Glen FC, Crabb DP, Smith ND, Burton R, Garway-Heath DF. 2012. Do patients with glaucoma have difficulty recognizing faces? *Invest Ophthalmol Vis Sci*. 53:3629–3637.
Goffaux V. 2016. Book review: perception of pixelated images. *Front Psychol*. 7:1–3.
Goffaux V, Peters J, Haubrechts J, Schiltz C, Jansma B, Goebel R. 2011. From coarse to fine? Spatial and temporal dynamics of cortical face processing. *Cereb Cortex*. 21:467–476.
Goffaux V, Rossion B. 2006. Faces are "spatial" - holistic face perception is supported by low spatial frequencies. *J Exp Psychol Hum Percept Perform*. 32:1023–1039.
Gold J, Bennett PJ, Sekuler AB. 1999. Identification of band-pass filtered letters and faces by human and ideal observers. *Vision Res*. 39:3537–3560.
Groner MT, Groner R, Von Mühlenen A. 2008. The effect of spatial frequency content on parameters of eye movements. *Psychol Res*. 72:601–608.
Guyader N, Chauvin A, Boucart M, Peyrin C. 2017. Do low spatial frequencies explain the extremely fast saccades towards human faces? *Vision Res*. 133:100–111.
Hacker CM, Meschke EX, Biederman I. 2019. A face in a (temporal) crowd. *Vision Res*. 157:55–60.
Hamilton R, Bach M, Heinrich SP, Hoffmann MB, Odom JV, McCulloch DL, Thompson DA. 2021. VEP estimation of visual acuity: a systematic review. *Doc Ophthalmol*. 142:25–74.
Harmon LD. 1973. The recognition of faces. *Sci Am*. 229:70–83.
Harmon LD, Julesz B. 1973. Masking in visual recognition: effects of two-dimensional filtered noise. *Science*. 180: 1194–1197.
Hayes T, Morrone MC, Burr DC. 1986. Recognition of positive and negative bandpass-filtered images. *Perception*. 15:595–602.
Hegdé J. 2008. Time course of visual perception: coarse-to-fine processing and beyond. *Prog Neurobiol*. 84:405–439.
Hemptinne C, Liu-Shuang J, Yuksel D, Rossion B. 2018. Rapid objective assessment of contrast sensitivity and visual

acuity with sweep visual evoked potentials and an extended electrode array. *Invest Ophthalmol Vis Sci*. 59:1144–1157.
Hole GJ, George PA, Eaves K, Rasek A. 2002. Effects of geometric distortions on face-recognition performance. *Perception*. 31:1221–1240.
Hsiao JH, Cottrell G. 2008. Two fixations suffice in face recognition. *Psychol Sci*. 19:998–1006.
Hsiao FJ, Hsieh JC, Lin YY, Chang Y. 2005. The effects of face spatial frequencies on cortical processing revealed by magnetoencephalography. *Neurosci Lett*. 380:54–59.
Jacobs DS, Blakemore C. 1988. Factors limiting the postnatal development of visual acuity in the monkey. *Vision Res*. 28:947–958.
Johnston RA, Edmonds AJ. 2009. Familiar and unfamiliar face recognition: a review. *Memory*. 17:577–596.
Keil MS. 2008. Does face image statistics predict a preferred spatial frequency for human face processing? *Proc R Soc B: Biol Sci*. 275:2095–2100.
Lane J, Rohan EMF, Sabeti F, Essex RW, Maddess T, Barnes N, He X, Robbins RA, Gradden T, McKone E. 2018a. Improving face identity perception in age-related macular degeneration via caricaturing. *Sci Rep*. 8:1–10.
Lane J, Rohan EMF, Sabeti F, Essex RW, Maddess T, Dawel A, Robbins RA, Barnes N, He X, McKone E. 2018b. Impacts of impaired face perception on social interactions and quality of life in age-related macular degeneration: a qualitative study and new community resources. *PLoS One*. 13:e0209218.
Le Grand R, Mondloch CJ, Maurer D, Brent HP. 2003. Expert face processing requires visual input to the right hemisphere during infancy. *Nat Neurosci*. 6:1108–1112.
Lee Y, Matsumiya K, Wilson HR. 2006. Size-invariant but viewpoint-dependent representation of faces. *Vision Res*. 46:1901–1910.
Loftus GR, Harley EM. 2005. Why is it easier to identify someone close than far away? *Psychon Bull Rev*. 12:43–65.
Millward R, O'Toole A. 1986. Recognition memory transfer between spatial-frequency analyzed faces. In: *Aspects of face processing*. Dordrecht: Springer, pp. 34–44.
Morrison DJ, Schyns PG. 2001. Usage of spatial scales for the categorization of faces, objects, and scenes. *Psychon Bull Rev*. 8:454–469.
Mousavi SM, Oruc I. 2020. Size effects in the recognition of blurry faces. *Perception*. 49:222–231.
Näsänen R. 1999. Spatial frequency bandwidth used in the recognition of facial images. *Vision Res*. 39:3824–3833.
Ojanpää H, Näsänen R. 2003. Utilisation of spatial frequency information in face search. *Vision Res*. 43:2505–2515.
Oruç I, Barton JJS. 2010. Critical frequencies in the perception of letters, faces, and novel shapes: evidence for limited scale invariance for faces. *J Vis*. 10:1–12.
Owsley C, Sekuler R, Boldt C. 1981. Aging and low-contrast vision: face perception. *Invest Ophthalmol Vis Sci*. 21: 362–365.
Parker DM, Lishman JR, Hughes J. 1992. Temporal integration of spatially filtered visual images. *Perception*. 21:147–160.
Parker DM, Costen NP. 1999. One extreme or the other or perhaps the golden mean? Issues of spatial resolution in face processing. *Current Psychology*. 18:118–127.
Parker DM, Lishman JR, Hughes J. 1997. Evidence for the view that temporospatial integration in vision is temporally anisotropic. *Perception*. 26:1169–1180.
Peli E, Goldstein RB, Young GM, Trempe CL, Buzney SM. 1991. Image enhancement for the visually impaired: simulations and experimental results. *Invest Ophthalmol Vis Sci*. 32:2337–2350.
Peli E, Lee E, Trempe CL, Buzney S. 1994. Image enhancement for the visually impaired: the effects of enhancement on face recognition. *J Opt Soc Am A*. 11:1929.
Petras K, Ten Oever S, Jacobs C, Goffaux V. 2019. Coarse-to-fine information integration in human vision. *Neuro Image*. 186:103–112.
Quek GL, Liu-Shuang J, Goffaux V, Rossion B. 2018. Ultra-coarse, single-glance human face detection in a dynamic visual stream. *Neuro Image*. 176:465–476.
Ramon M, Vizioli L, Liu-Shuang J, Rossion B. 2015. Neural microgenesis of personally familiar face recognition. *Proc Natl Acad Sci USA*. 112:E4835–E4844.
Rao RPN, Ballard DH. 1999. Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nat Neurosci*. 2:79–87.
Regan D. 1973. Rapid objective refraction using evoked brain potentials. *Invest Ophthalmol*. 12:669–679.
Retter TL, Jiang F, Webster MA, Rossion B. 2020. All-or-none face categorization in the human brain. *Neuro Image*. 213:116685.
Retter TL, Rossion B. 2016. Uncovering the neural magnitude and spatio-temporal dynamics of natural image categorization in a fast visual stream. *Neuropsychologia*. 91:9–28.
Rossion B. 2013. The composite face illusion: a whole window into our understanding of holistic face perception. *Vis Cognit*. 21:139–253.
Rossion B. 2018. Humans are visual experts at unfamiliar face recognition. *Trends Cogn Sci*. 22:471–472.
Rubin G, Siegel K. 1984. Recognition of low-pass faces and letters (ARVO supplement). *Invest Ophthalmol Vis Sci*. 25:96.
Ruiz-Soler M, Beltran FS. 2006. Face perception: an integrative review of the role of spatial frequencies. *Psychol Res*. 70:273–292.
Schwaninger A, Lobmaier JS, Collishaw SM. 2002. Role of featural and configurai information in familiar and unfamiliar face recognition. *Lecture Note Comput Sci*. 2525:643–650.
Schyns PG, Bonnar L, Gosselin F. 2002. Show me the features! Understanding recognition from the use of visual information. *Psychol Sci*. 13:402–409.
Sergent J. 1986. Microgenesis of face perception. In: *Aspects of face processing*. Dordrecht: Springer, pp. 17–33.
Shahangian K, Oruc I. 2014. Looking at a blurry old family photo. *Perception*. 43:90–98.
Sinha P, Poggio T. 1996. I think I know that face. *Nature*. 384:404.
Sinha P. 2002. Identifying perceptually significant features for recognizing faces. In: *In Human vision and electronic imaging VII*. Vol 4662. San Jose, (CA): International Society for Optics and Photonics, pp. 12–21.
Strang G, Nguyen T. 1996. *Wavelets and filter banks*. Wellesley: Wellesley-Cambridge Press.
Tanaka JW, Farah MJ. 1993. Parts and wholes in face recognition. *Q J Exp Psychol Sec A*. 46:225–245.
Tanskanen T, Näsänen R, Montez T, Päällysaho J, Hari R. 2005. Face recognition and cortical responses show similar sensitivity to noise spatial frequency. *Cereb Cortex*. 15:526–534.
Tavassoli A, van der Linde I, Bovik AC, Cormack LK. 2009. Eye movements selective for spatial frequency and orientation during active visual search. *Vision Res*. 49:173–181.

Trapp S, Pascucci D, Chelazzi L. 2021. Predictive brain: addressing the level of representation by reviewing perceptual hysteresis. *Cortex*.

Vogelsang L, Gilad-Gutnick S, Ehrenberg E, Yonas A, Diamond S, Held R, Sinha P. 2018. Potential downside of high initial visual acuity. *Proc Natl Acad Sci USA*. 115:11333–11338.

Watier NN, Collin CA, Boutet I. 2010. Spatial-frequency thresholds for configural and featural discriminations in upright and inverted faces. *Perception*. 39:502–513.

Watt RJ. 1987. Scanning from coarse to fine spatial scales in the human visual system after the onset of a stimulus. *J Opt Soc Am A*. 4:2006.

Yan X, Rossion B. 2020. A robust neural familiar face recognition response in a dynamic periodic stream of unfamiliar faces. *Cortex*. 132:281–295.

Yan X, Young AW, Andrews TJ. 2017. The automaticity of face perception is influenced by familiarity. *Atten Percept Psychophys*. 79:2202–2211.

Yan X, Zimmermann FGS, Rossion B. 2020. An implicit neural familiar face identity recognition response across widely variable natural views in the human brain. *Cogn Neurosci*. 11:143–156.

Yin RK. 1969. Looking at upside-down faces. *J Exp Psychol*. 81:141.

Yip AW, Sinha P. 2002. Contribution of color to face recognition. *Perception*. 31:995–1003.

Young AW, Burton AM. 2017. Recognizing faces. *Curr Dir Psychol Sci*. 26:212–217.

Young AW, Hellawell D, Hay DC. 1987. Configural information in face perception. *Perception*. 16:747–759.

Zhao L, Chubb C. 2001. The size-tuning of the face-distortion after-effect. *Vision Res*. 41:2979–2994.

Zimmermann FGS, Yan X, Rossion B. 2019. An objective, sensitive and ecologically valid neural measure of rapid human individual face recognition. *R Soc Open Sci*. 6:181904.