Neuropsychologia 104 (2017) 182–200

Contents lists available at ScienceDirect

Neuropsychologia

journal homepage: www.elsevier.com/locate/neuropsychologia

Category-selective human brain processes elicited in fast periodic visual stimulation streams are immune to temporal predictability

Genevieve L. Quek[a], Bruno Rossion[a,b,*]

[a] Psychological Sciences Research Institute, Institute of Neuroscience, University of Louvain, Belgium
[b] Neurology Unit, Center Hospitalier Regional Universitaire (CHRU) de Nancy, F-54000 Nancy, France

ARTICLE INFO

Keywords:
Periodicity
Perceptual categorization
Faces
Neurophysiology

ABSTRACT

Recording direct neural activity when periodically inserting exemplars of a particular category in a rapid visual stream of other objects offers an objective and efficient way to quantify perceptual categorization and characterize its spatiotemporal dynamics. However, since periodicity entails predictability, perceptual categorization processes identified within this framework may be partly generated or modulated by temporal expectations. Here we present a stringent test of the hypothesis that temporal predictability generates or modulates category-selective neural processes as measured in a rapid periodic visual stimulation stream. In Experiment 1, we compare neurophysiological responses to periodic and nonperiodic (i.e., unpredictable) variable face stimuli in a fast (12 Hz) visual stream of nonface objects. In Experiment 2, we assess potential responses to rare (10%) omissions of periodic face events (i.e., violations of periodicity) in the same fast visual stream. Overall, our observations indicate that category(face)-selective processes elicited in a fast periodic stream of visual objects are immune to temporal predictability. These observations do not support a predictive coding framework interpretation of category-change detection in the human brain and have important implications for understanding automatic human perceptual categorization in a rapidly changing (i.e., dynamic) visual scene.

1. Introduction

The human brain is able to rapidly and effortlessly organize visual information in the environment. With just a single glance, we can tell almost instantly that a roundish object in our field of view is a face – not a flower, a tennis ball, a clock, or any other type of object. This ability to rapidly group currently experienced stimuli into meaningful categories – known as perceptual categorization – is surely one of the most fundamental high-level brain functions, serving as the foundation for memory, learning, language, affective processing, decision making, and action execution.

In the visual domain, a powerful way to shed light on perceptual categorization processes is to combine visual periodicity with direct recording of neural activity, for instance using electroencephalography (EEG). By embedding members of a specific category at a strict periodic rate within a dynamic visual stream of items that do *not* belong to that category, perceptual categorization processes of interest are projected to a specified frequency in the EEG spectrum. At a rapid (and quasi-continuous) rate, this approach can isolate category-selective visual processes without post-hoc subtraction, in a manner that is both objective and highly efficient (Jacques et al., 2016; Jonas and Rossion, 2016; Retter and Rossion, 2016). For example, Lochy et al. (2015) investigated lexical categorization processes by presenting participants with a stream of non-word items a rate of exactly 10 Hz (i.e., 10 non-words per second), with a word stimulus embedded as every fifth item. Three minutes of this stimulation elicited an electrophysiological response at the exact frequency of image presentation (i.e., 10 Hz), but more importantly, a robust response at the exact periodicity of the word items embedded in non-word sequence (i.e., 10 Hz/5 items = 2 Hz), even in the absence of an overt lexical decision task. The authors interpreted this 2 Hz signal to be a differential response to words compared to non-words, as it could only have arisen if the response evoked by words *differed* from that evoked by non-words (see also Lochy et al., 2016). The same periodicity-based approach (i.e., Fast Periodic Visual Stimulation, or FPVS) has also been used to examine human adults and infants' perceptual categorization of faces and natural object images (e.g., Fig. 1; de Heering and Rossion, 2015; Rossion et al., 2015; Jacques et al., 2016; Retter and Rossion, 2016). For example, Retter and Rossion (2016) presented participants with a dynamic stream of object images at a rate of 12.5 Hz (i.e., 80 ms per image), inserting face images

* Correspondence to: Institute of Research in Psychology (IPSY) & Institute of Neuroscience (IoNS), University of Louvain 10 Place du Cardinal Mercier, Louvain-la-Neuve 1348, Belgium.
E-mail address: bruno.rossion@uclouvain.be (B. Rossion).

http://dx.doi.org/10.1016/j.neuropsychologia.2017.08.010
Received 15 March 2017; Received in revised form 12 July 2017; Accepted 5 August 2017
Available online 12 August 2017
0028-3932/

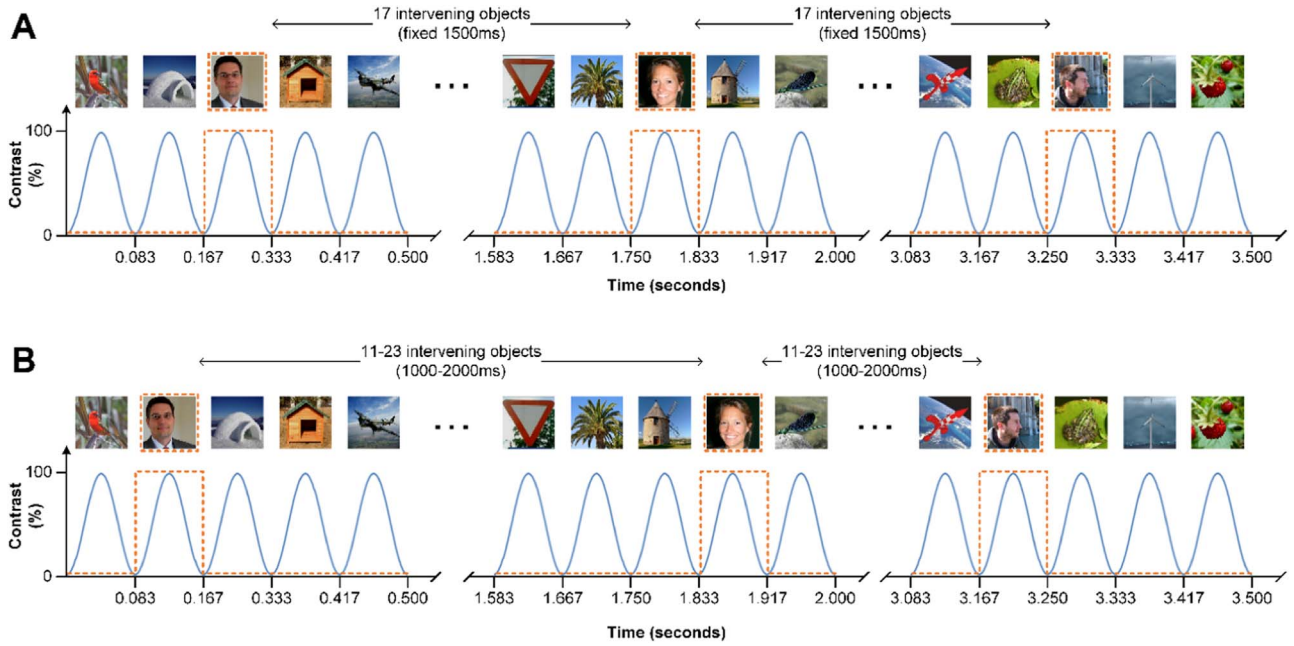


Fig. 1. Expt. 1 sequence design. We presented images at a rate of exactly 12 Hz by sinusoidally modulating the contrast of each from 0–100-0% (blue solid lines). In both conditions, the 90 s sequence contained 60 natural face images and numerous natural object images (e.g., vehicles, animals, buildings, trees, etc.). (A) In the periodic condition, faces appeared at regular intervals every 18 stimuli (orange dashed lines). (B) In the nonperiodic condition, faces were spaced at irregular intervals, appearing anywhere after 11–23 object images (orange dashed lines). Participants did not respond to the faces, but instead monitored a central fixation cross overlaid on the images for color changes. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article).

into the sequence every three, five, seven, nine, or 11 stimuli. In addition to finding a strong response at the image presentation frequency of 12.5 Hz, they also observed a robust category-selective response at each of the defined face periodicities (e.g., a face every seven images gives a response at exactly 12.5 Hz/7 = 1.79 Hz), indicating there was a differential response to faces as compared to objects. Given the very rapid image presentation rate (each image was replaced after just 80 ms), and the use of a completely orthogonal task, the authors argued that this category-selective response reflected automatic categorization of faces vs. objects at the perceptual rather than decisional level. This conclusion is supported by the application of this approach to intracerebral recordings in a large group of human patients, identifying and quantifying the face-selective responses in localized regions of the right ventral occipito-temporal cortex (Jonas and Rossion, 2016).

Since the response of interest in FPVS designs depends on *i)* the critical stimuli and the temporally surrounding distractors evoking different responses, and *ii)* measuring a similar evoked response each time a critical stimulus appears, using a large number of highly variable exemplars (e.g., 50 natural face images and 200 natural object images in Rossion et al. (2015)) ensures this response will capture both the degree to which the visual system is able to *discriminate* the critical category from others in the stream, as well as the extent to which it is able to *generalize* across widely differing exemplars within the category (Jonas and Rossion, 2016; Rossion et al., 2015; Retter and Rossion, 2016). Importantly, the reliance here on a *periodic* response also serves to minimize low-level image confounds without artificially standardizing low-level stimulus properties. When highly variable natural images are used, the amplitude spectra of two categories may vary *on average*, but will not vary consistently across a stimulus set. As such, a given set of low-level cues will not occur reliably at the critical category frequency, where the response of interest is measured. This claim is borne out by the observation that phase-scrambled natural images, in which the amplitude spectra are preserved, but structural information is removed, do not elicit category-selective responses in FPVS designs (de Heering and Rossion, 2015; Rossion et al., 2015; for an extended discussion see Retter and Rossion, 2016)

An objective measure of high-level category-selective processing that taps both between-category discrimination and within-category generalization – i.e., the core abilities which underlie successful perceptual categorization in natural settings – is an exciting development for the field of visual perception. Moreover, the high signal-to-noise ratio enjoyed by the approach makes it an ideal method for testing young children or clinical populations, who may have particularly noisy EEG signals. Yet an important and outstanding theoretical issue is whether the category-selective signal yielded by periodicity is generated in part by temporal expectation. That is, since the critical category exemplars always appear at periodic intervals in FPVS, and since the entire sequence is itself a rhythmic stimulation (Jones, 1976), participants in these tasks could conceivably form reliable expectations (either explicit or implicit) about exactly *when* critical stimuli will appear (McAuley and Jones, 2003). Indeed, a number of studies have shown that regular ("rhythmic") stimulation can induce strong temporal expectations, thereby facilitating sensory processing of stimuli both in the auditory (e.g., Morillon et al., 2016) and visual domains (Mathewson et al., 2010; Rohenkohl et al., 2012; Cravo et al., 2013; Breska and Deouell, 2014). Generally, these effects are expressed in terms of greater encoding precision, higher perceptual sensitivity and decreased response times in behavioral tasks (Rajendran and Teki, 2016). Moreover, behavioral studies employing rapid serial visual presentation (RSVP; Potter and Levy, 1969), a stimulation that is similar in kind to FPVS, have shown that identification accuracy for targets embedded in these streams improves as a function of number of distractors before target onset, suggesting that temporal expectation is "tuned" over the course of the RSVP sequence itself (Ariga and Yokosawa, 2008).

If the rhythmicity of the FPVS approach (e.g., images appearing at a defined periodic rate), combined with the temporal predictability of critical category exemplars (e.g., a face after every 9 object images), does indeed elicit temporal expectations in participants, then the category-selective response it yields may not solely reflect processes related to perceptual categorization, but may be generated in part by temporal expectation. As a case in point, the category-selective response for faces embedded in a stream of objects is known to be comprised of several components starting at ~ 100 ms and lasting up to ~ 500 ms after face onset (Rossion et al., 2015; Jacques et al., 2016;

Retter and Rossion, 2016). Given that behavioral detection of a face can be achieved within 100–300 ms following stimulus onset (Rousselet et al., 2003; Crouzet et al., 2010; Crouzet and Thorpe, 2011), it could well be the case that the later aspects of this neural response may be generated or at least largely modulated by temporal expectation about the appearance of faces in the visual stream. Indeed, faces are perhaps the ideal stimulus with which to test the role of temporal expectation in driving the category-selective response in periodicity-based designs, since observers can easily form a visual template for this category based on the cardinal arrangement of facial features. Moreover, face perception is known to be highly sensitive to top-down effects, as evidenced by phenomena like the hollow-face illusion (Gregory, 1970) and face recognition of 2-tone pictures ("Mooney faces"; Mooney, 1957; Cavanagh and Leclerc, 1989; Moore and Cavanagh, 1998).

Here we present two EEG experiments that shed light on the important and unresolved issue of whether temporal predictability modulates implicit perceptual categorization in the context of dynamic visual stimulation. In Experiment 1, we compare the category-selective response elicited by FPVS sequences in which the critical face stimuli are either temporally predictable (embedded at periodic intervals) or unpredictable (embedded at nonperiodic intervals). If temporal expectations about faces contribute to the category-selective response, then the response to periodic and nonperiodic faces should reliably differ. In Experiment 2, we examine the neural response to a violation of temporal predictability in an FPVS sequence, by introducing rare random omissions of periodic (i.e., expected) faces after a period of familiarization. Here we predict that non-face stimuli which violate temporal expectations (i.e., missing faces) should elicit a different response than non-face stimuli about which participants cannot build any temporal expectations.

2. Experiment 1

Our goal in Experiment 1 was to dissociate the relative contribution of two factors that are typically conflated in FPVS designs – periodicity and temporal predictability. To this end, we embedded face stimuli at either fixed or random intervals in a sequence of variable object images. After verifying this manipulation was effective, we employed a novel approach we term 'False-Sequencing' to impose a new face periodicity of exactly 1 Hz on both the predictable (i.e., periodic) and unpredictable (i.e., nonperiodic) conditions. In this analysis, face periodicity is held constant (always 1 Hz) while temporal predictability is allowed to vary. If temporal predictability confers a processing advantage on the critical face stimuli in FPVS sequences, then the category-selective response elicited by temporally predictable and unpredictable faces will differ in some way. Under a predictive coding framework, we might expect this difference to take the form of an attenuated response to periodic faces as compared to nonperiodic faces, since sensory input corresponding to the former stimulus is consistent with current high level expectations and thus effectively redundant (Rao and Ballard, 1999; Friston, 2005; Alink et al., 2010; Kok et al., 2012a). On the other hand, if temporal expectation sharpens sensory representation by suppressing neural responses that are inconsistent with current expectations (Lee and Mumford, 2003), then periodic faces might be expected to elicit a stronger category-selective response than nonperiodic faces, since noise is reduced in the former case.

3. Experiment 1 methods

3.1. Participants

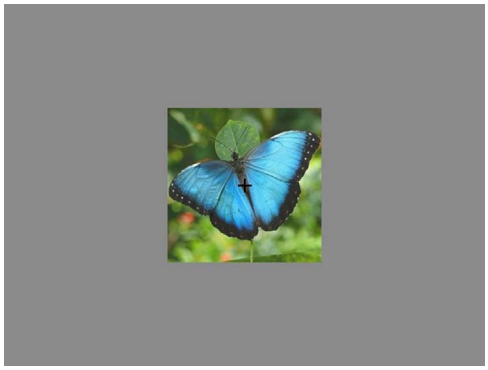
Twenty persons (nine females) aged between 18 and 26 years took part in this study in exchange for payment. All were right-handed, with normal or corrected-to-normal vision, and had no history of neurological illness. In accordance with the University of Louvain Biomedical Ethics Committee guidelines, we obtained written informed consent from all participants prior to testing.

3.2. Stimuli

Stimuli were 200 color images of various non-face objects (e.g., animals, plants, manmade structures/objects) and 100 color images of human faces. We resized each image to 256 ×256 pixels and equalized their mean luminance values in MatLab (R2012b). As in our previous studies that used a different but comparable stimulus set (e.g., Rossion et al., 2015), we deliberately did not segment faces/objects from their original naturalistic backgrounds, such that the viewpoint, lighting, and image composition varied widely across the full stimulus set (see Fig. 1 for examples). During the stimulation sequence, each image appeared on a uniform grey background and subtended 9.077° of visual angle. A small black fixation cross (subtending 0.788° of visual angle) overlaid the images throughout the sequence.

3.3. Procedure

Participants sat in a darkened room and viewed a computer monitor at a distance of 80 cm. They viewed stimulation sequences consisting of a two second pre-stimulation period with only the central fixation cross, a two second fade-in, a 90 s sequence of rapid images presented at a periodic rate of exactly 12 Hz (12 images per second), and a three second fade-out. We used a 12 Hz stimulation rate since generic face-categorization (i.e., the detection of faces in the visual environment) can be achieved within the 83 ms image duration that 12 Hz provides (e.g., Retter and Rossion, 2016). Second, as 12 Hz is well above the range of frequencies that leads to large face-selective responses (i.e., 6 Hz, Alonso-Prieto et al., 2013), this image presentation rate allows us to better dissociate face-selective processing from general visual processing. We used customized software programmed in Java to sinusoidally modulate the contrast of each image from 0% to 100–0% over a period of 83.33 ms (see Fig. 1; Movie 1). Since stimuli are still visible at low contrast, this ensured a near-continuous stimulation, with no perceptual interruption between images (see Retter and Rossion, 2016, for a comparison of sinusoid and squarewave stimulation). We instructed participants to fix their gaze on the central cross overlaid on the images, and to press a key whenever it changed color (200 ms change duration; 12 changes in each stimulation sequence; minimum 2 s (s) between consecutive changes). In the periodic condition, sequences contained randomly selected non-face images with a face image embedded as every 18th stimulus (Fig. 1A). This stimulation should elicit two periodic EEG responses: one at 12 Hz, reflecting visual processing common to both face and non-face stimuli (referred to as the common response), and one at 12 Hz/18 (i.e., 0.67 Hz), reflecting the differential response to faces compared to other images (referred to as the category-selective or face-selective response, Retter and Rossion, 2016). Nonperiodic sequences were identical to periodic sequences, save that the face images appeared at *irregular* intervals during the rapid sequence (11–23



Video S1. Example stimulation sequence from periodic condition (90 s), in which natural object images appeared at a fast periodic rate of exactly 12 Hz, with various natural face images embedded every 18 stimuli (i.e., 12 Hz/18 = 0.67 Hz)". Supplementary material related to this article can be found online at http://dx.doi.org/10.1016/j.neuropsychologia.2017.08.010.

non-face images between the face images, see Fig. 1B). Here we expect to find only the common response at 12 Hz, since there is no second periodicity embedded in the sequence. Participants viewed four stimulation sequences of each type in a pseudorandom order, each containing exactly 60 face image presentations.

3.4. EEG acquisition

We acquired high-density 128-channel EEG during sequence viewing using an ActiveTwo Biosemi system with a 512 Hz sampling rate (http://www.biosemi.com/). We monitored eye-movements via four external electrodes placed at the outer canthi of the eyes, and above and below the right eye. The electrode offset criterion cut-off was 20 μV.

3.5. Analysis

3.5.1. Behavioral analysis

We calculated the accuracy and response time (RT) for each fixation cross color change during the stimulation sequence, removing values that occurred < 250 ms or > 1500 ms following change onset.

3.5.2. EEG analysis

We analyzed EEG data using the open source toolbox LetsWave5 (http://nocions.webnode.com/letswave) running on MatLab R2012b (MathWorks, USA). Filtering was implemented using the *'butter'* and *'filtfilt'* functions of this toolbox. All EEG waveform analysis procedures described here have been extensively documented in recent publications (Liu-Shuang et al., 2014; Rossion, 2014; Rossion et al., 2015; Jacques et al., 2016; Retter and Rossion, 2016). We carried out all statistical analyses in R (https://www.r-project.org/), using repeated measures analysis of variance (ANOVA) to examine main effects and interactions. We used Greenhouse-Geisser corrections to adjust degrees of freedom wherever the assumption of sphericity was violated. Significant main effects were followed up with posthoc Tukey HSD contrasts, and significant interactions decomposed using two-tailed paired *t*-tests (Bonferroni corrected).

3.5.2.1. EEG pre-processing. We excluded data from two male participants due to excessive artefact on multiple channels in the EEG trace. Before removing these participants, we verified that both showed a significant face-selective response in the EEG spectrum. For each of the remaining 18 participants, we removed slow voltage changes of non-neural origin and irrelevant EMG artefact by applying a zero phase-shift Butterworth band-pass filter (4th order, 0.05–100 Hz) to the raw EEG trace (Luck, 2005). We then removed AC electrical noise at 50 Hz, 100 Hz, and 150 Hz using an FFT multi-notch filter (Hanning window, width = 0.5 Hz), and then segmented 99 s epochs relative to the starting trigger of each stimulation sequence (−2 s to 97 s). To remove blink artefact, we applied an independent components analysis (ICA) with a square mixing matrix (Hyvarinen and Oja, 2000) and removed a single component identified by visual inspection of the waveform and corresponding topography. Artefact-prone channels with deflections greater than 100 μV in at least two trials were replaced using linear interpolation of neighboring clean channels (less than 5% of channels per participant). Since we recorded high-density EEG, we re-referenced each channel's signal using the mean of all 128 scalp channels (rather than single reference electrode), and relabeled each to match the standard 10–20 system (for additional detail see Rossion et al., 2015).

3.5.2.2. Frequency-domain analysis. Following pre-processing, we segmented an epoch for each sequence containing an integer number of cycles of the face presentation frequency (0.67 Hz), resulting in an 88.51 s epoch containing exactly 59 face presentation cycles for each stimulation sequence. This ensured that a single frequency bin would be centered on the category-selective frequency of interest (0.67 Hz) in the EEG spectrum. We averaged the four epochs for each condition before applying a Fast Fourier Transformation (FFT) to extract a normalized amplitude spectrum for each channel ranging between 0 and 256 Hz. The frequency resolution of these spectra was very high, as determined by the inverse of the sequence duration (i.e., 1/88.5 = 0.0113 Hz). To identify significant response signal at the relevant stimulation frequencies (i.e., 12 Hz, 0.67 Hz, and harmonics), we computed a *z*-score at each discrete frequency bin (e.g., Liu-Shuang et al., 2014; Rossion, 2014; Rossion et al., 2015; Jacques et al., 2016; Retter and Rossion, 2016) using the amplitude spectra pooled across all participants, conditions, and scalp channels. Here we specified a noise range comprised of the 20 frequency bins neighboring the frequency of interest, excluding the immediately adjacent bins and the local maximum and minimum amplitude bins (e.g., Rossion and Boremanse, 2011; Rossion et al., 2012; Retter and Rossion, 2016). We subtracted the mean amplitude of this noise range from the amplitude at the frequency of interest, and divided the result by the standard deviation of amplitudes in the noise range. The advantage of determining statistical significance this way is that unlike *t*-tests (which rely in inter-individual variance), z scores can computed either at the group level or for each participant individually (for an extended discussion of detecting significant signal at a specified frequency, see Appendix 2 of Norcia et al., 2015). As in our previous studies (e.g., Jacques et al., 2016), we considered response signals with a *z*-score greater than 3.1 to be significant ($p < .001$, one tailed, i.e., signal > noise), and selected only these frequencies for further analysis. Note that the use of one-tailed tests is well justified in this context, as the goal here is to identify frequencies at which the signal is significantly *greater* than the noise in the surrounding bins (instances in which the signal is significantly *lower* than the surrounding bins are of no relevance).

To take into account the variation in noise across the EEG spectrum, we performed a local baseline-subtraction on the raw amplitudes for each condition using the same baseline range as for z-score calculation (see above). To quantify the category-selective response, we summed the baseline-corrected values across all relevant harmonic frequencies (Jacques et al., 2016; Retter and Rossion, 2016) in two ways: *i)* at the global level by pooling the information across all scalp channels, and *ii)* within functional regions-of-interest (ROIs) compatible with the stable bilateral occipito-temporal pattern typically elicited by FPVS paradigms (Rossion et al., 2015; Jacques et al., 2016; Retter and Rossion, 2016). We defined the ROIs by averaging the four channels in each hemisphere with the greatest summed-harmonic response averaged across-conditions, resulting in a left occipito-temporal ROI (averaged over channels P9, PO7, PO9, and PO11) and a right occipito-temporal ROI (averaged over channels P10, PO8, PO10, and PO12). The central occipito-parietal ROI was the average of the four channels with the largest common response at 12 Hz (averaged over channels Oz, POOz, PO4h, and POO6).

3.5.2.3. False-sequencing analysis. To dissociate the contribution of periodicity and temporal predictability to the category-selective response, we used a novel 'false-sequencing' approach. For each stimulation sequence in each condition, we segmented a 1000 ms epoch around each face onset (−500 ms to 500 ms), resulting in ~60 segments per sequence. We sequenced these 1000 ms segments by aligning the first sample of each epoch with the last sample of the preceding epoch, producing a new 60 s continuous EEG trace. To compensate for introduced drift in this newly created 'false-sequence', we subtracted the mean voltage on each channel and linearly detrended the data. We then cropped each false-sequence to be exactly 55 s (corresponding to an integer number of number of cycles of 0.67 Hz), and averaged these by condition. Remaining analyses for the false-sequence data were as described for the frequency-domain analysis above. Additionally, we used the 'BayesFactor' package in R (http://bayesfactorpcl.r-forge.r-project.org/) to calculate a Bayes factor for each possible model of the false-sequence data containing the factors *Periodicity* and *ROI*.

3.5.2.4. Time-domain analysis. In a separate analysis, we compared the response to periodic and nonperiodic faces in the time-domain. Here we applied a zero phase-shift Butterworth low-pass filter (4th order, 30 Hz cut off) to the raw EEG trace before cropping each stimulation sequence to an integer number of cycles of the image presentation rate (i.e., 12 Hz). We then isolated the differential response to faces compared to objects, by applying an FFT multi-notch filter (Hanning window, 0.05 Hz width) at 12 Hz, 24 Hz, and 36 Hz to selectively remove signal corresponding to the image presentation rate. We then segmented a 1000 ms epoch around each face onset (−200 ms to 800 ms) and averaged the resulting ~ 240 epochs for each condition. We applied a standard baseline-correction procedure by subtracting the mean amplitude in the −167 to 0 ms time window preceding face onset (i.e., 2 cycles at 12 Hz). In each ROI, we compared the conditional mean waveforms by conducting a two-tailed paired *t*-test at each of the 513 time points. We inspected these obtained *t*-values relative to two significance criteria ($p < .01$ uncorrected, and $p < .05$, corrected for multiple comparisons). We corrected for multiple comparisons using a permutation procedure capable of maintaining experiment-wise error in time-series analysis (Blair and Karniski,1993; Finkbeiner and Friedman, 2011). The steps for this procedure were as follows. First we generated 100,000 permutations of our data by systematically shuffling the periodic/nonperiodic labels for each participant's conditional mean datasets (consisting of 513 time points). Assuming the null hypothesis is true (i.e., there is no difference between periodic and nonperiodic waveforms), the assignment of condition labels should be arbitrary, such that each instance of permutated data is just as likely to have occurred as our actual obtained data. In this way, the permuted data arrangements represent 100,000 possible outcomes of our experiment. For each permutation, we conducted a two-tailed paired *t*-test at every time point ($n = 513$, for a total of 51.3 million comparisons) and kept aside the maximum *t*-value ($t_{|max|}$) given by that permutation. Rank ordering the $t_{|max|}$ values obtained across all permutations forms a reference distribution against which each obtained *t* value can be compared. To maintain experiment-wise α at 0.05, we took the value in the $t_{|max|}$ reference distribution that cut off 0.025 of each tail as our critical *t* (see Blair and Karniski, 1993 for a more detailed explanation of the rationale behind this procedure).

We assessed the similarity/difference of the time-domain response to periodic and nonperiodic faces in a second way using multivariate pattern classification (MVPA). Unlike the permutation testing described above, MVPA does not rely on single channel waveform data (i.e., a single averaged channel for each ROI), but considers fine differences in the pattern of activity elicited by periodic and nonperiodic faces across all channels simultaneously. We used an MPVA procedure for spatio-temporal decoding in EEG introduced by Bode and colleagues (The Decision Decoding Toolbox, available at http://ddtbox.github.io/DDTBOX/; Bode and Stahl, 2014; Bode et al., 2016). Using the Butterworth low-pass filtered data in which the image presentation frequency of 12 Hz was preserved (i.e., not yet selectively removed by notch-filtering), we extracted epochs corresponding to periodic faces and nonperiodic faces (−166.66 ms to 500 ms) for each participant. We used non-overlapping spatiotemporal analysis time-windows of 20 ms (10 data points) for each of the 128 channels. This time-window moved forward by 20 ms increments to cover the entire 666.66 ms epoch (Bai et al., 2007; Das et al., 2010; Blankertz et al., 2011; Bode et al., 2012; Bode and Stahl, 2014). At each time-window, we transformed the data into a spatiotemporal pattern vector labelled by face periodicity. We then trained an SVM classifier on 90% of the periodic and nonperiodic vectors using LIBSVM (Chang and Lin, 2011), and tested it on the remaining 10% of vectors. We repeated this classification process using a 10-fold cross-validation procedure which took different, non-overlapping chunks of test and train data on each pass. To further guard against any drawing biases that could influence classification, we repeated the ten cross-validation steps five times over, each time with a newly drawn 10% test portion of the data. On each pass, we calculated classification accuracy at each time-window, and then averaged across the 50 analyses to produce a final classification accuracy series for each subject. We further generated an empirical chance distribution generated using the exact same classification procedure, only with randomly shuffled labels for each iteration (i.e., each face epoch randomly labelled as either periodic or nonperiodic). Collated across participants, classification accuracies for the permutated data form a reference distribution at each time point, the mean of which we compared to the real classification accuracy using a two-tailed paired *t*-test (34 time points, significance criterion = $p < .05$, Bonferroni corrected). This method is considered stricter than testing against theoretical chance accuracy of 50% (Bode and Stahl, 2014; Bode et al., 2016).

In a complementary analysis, we used MVPA to ask whether the temporal unfolding of activity across the scalp elicited by a periodic face would *generalize* well to the activity elicited by a nonperiodic face. Again using pre-notch filtered data (i.e., 12 Hz response preserved), here we segmented an equal number of face and object instances for each periodicity condition and participant (−166.66 ms to 500 ms). We ensured an equal number of both stimulus types by taking only those object instances appearing exactly 6 cycles before each face. Using the same classification parameters as described above, we performed i) a *within-condition* decoding analysis of faces vs. objects by (classifier trained using 90% of face and object epochs from the *periodic* condition, and tested it on the remaining 10% of epochs from that same condition), and ii) a *cross-condition* decoding analysis (classifier on 90% of the *periodic* data, and tested it on 10% of the *nonperiodic* data). The rationale for cross-condition decoding is as follows – since the classifier is trained to distinguish neural activity evoked by an object from that evoked by a temporally predictable face, if it can perform at above-chance level in classifying neural activity elicited by objects and temporally unpredictable faces, then the response evoked by faces under predictable and unpredictable conditions must be similar. Alternatively, if the classifier trained to distinguish periodic face and object responses cannot generalize to nonperiodic face and object data, then we may conclude that response evoked by temporally regular and irregular faces is qualitatively different. In a final step, we compared the decoding accuracy for these two analyses at each time point using two-tailed paired *t*-tests (34 time points, significance criterion: $p < .05$, Bonferroni corrected). If cross-condition decoding (i.e., generalizing from periodic data to nonperiodic data) is significantly worse than within-condition decoding (i.e., generalizing from periodic data to periodic data), we may conclude that face periodicity changes the nature of the face response in FPVS paradigms.

4. Experiment 1 results

4.1. Behavior

Accuracy for the fixation cross color change task was uniformly high in both conditions, with no significant difference in percent correct between periodic trials ($M = 95.95\%$, $SD = 8.44\%$) and nonperiodic trials ($M = 96.41\%$, $SD = 6.46\%$), $t(17) = 0.403$, $p = 0.692$, $d = 0.09$. There was also no evidence that mean response time (RT) differed between the periodic ($M = 483$ ms, SD = 51 ms) and nonperiodic conditions ($M = 485$ ms, $SD = 53$ ms), $t(17) = 0.648$, $p = 0.526$, $d = 0.15$.

4.2. Frequency-domain results

4.2.1. Response at the category-selective frequency (0.67 Hz)

To avoid condition or channel related biases, we determined the range of frequencies for quantification of the category-selective response by inspecting the amplitude spectrum averaged across all conditions and all channels. Our z-score inspection procedure revealed highly significant responses at each of the first 13 consecutive harmonics of 0.67 Hz. When split by condition, all 13 of these harmonics

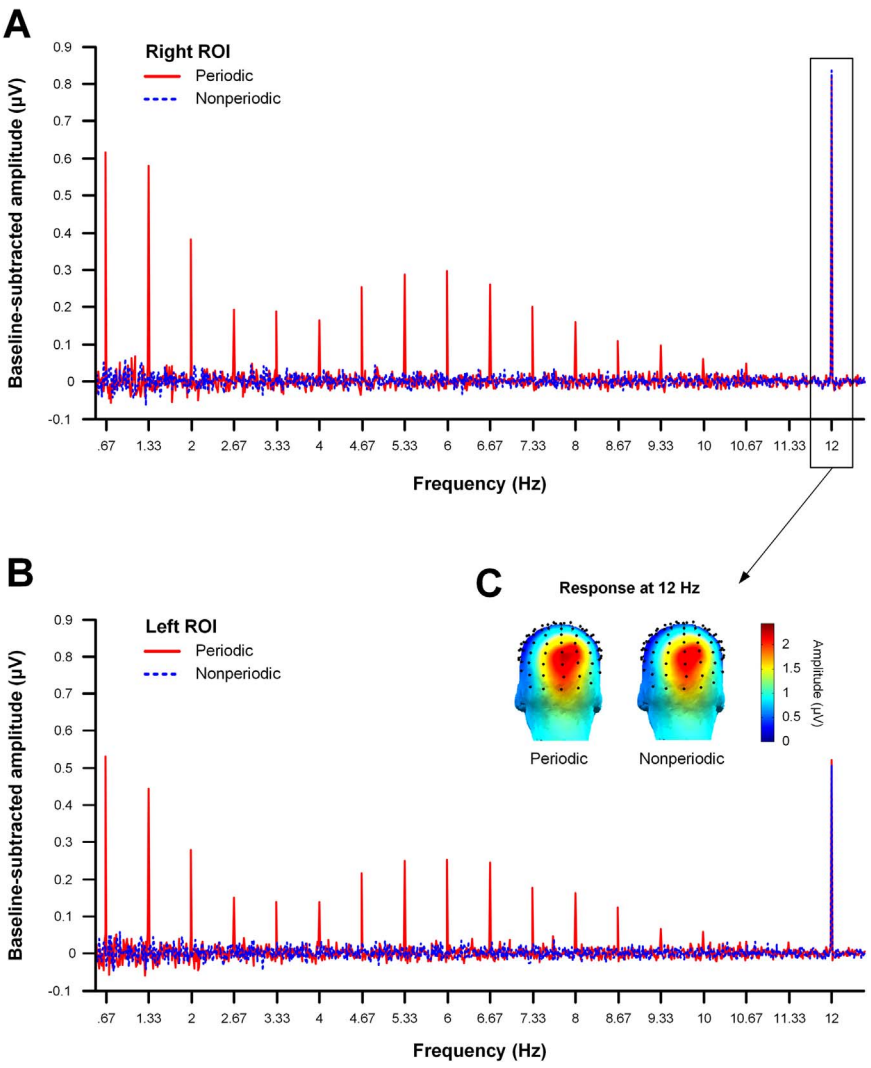


Fig. 2. Baseline-corrected amplitude spectra for Expt. 1, shown as a function of face periodicity for the **(A)** right and **(B)** left ROIs. In both regions, the response at the first 13 harmonics of the category-selective frequency (0.67 Hz) was significant ($p < .001$, one-tailed) in the periodic condition (red lines). In contrast, there was no response at any of these harmonics in the nonperiodic condition for either ROI. Importantly, both conditions yielded a significant response at the common frequency (i.e., 12 Hz) in both lateral ROIs. **(C)** This response was predominantly centered over medial occipital sites (around Oz).

remained significant in the periodic condition, but did not meet criterion in the nonperiodic condition. Accordingly, the sum of baseline-corrected amplitudes across these 13 harmonics was much larger in the periodic condition ($M = 1.162$ μV, $SD = 0.466$ μV) compared to the nonperiodic condition ($M = 0.038$ μV, $SD = 0.084$ μV), $t(17) = -9.64$, $p < .0001$, $d = 2.27$. This same pattern held when we inspected the category-selective response as a function of ROI, in that each region had a strong category-selective response in the periodic condition, but not the nonperiodic condition (Fig. 2).

We quantified the category-selective response in each ROI/condition by summing the baseline-corrected amplitude values across the first 13 harmonics of 0.67 Hz. We subjected these aggregate values to a repeated measures ANOVA with the factors *Periodicity* and *ROI*. Here we observed a significant interaction, $F(2,34) = 23.86$, $p < .0001$, $\eta_G^2 = 0.19$, the nature of which is clear in Fig. 3. Where *ROI* clearly modulated the category-selective response in the periodic condition, it had no such impact in the nonperiodic condition, since the responses in this condition were at floor level. Follow-up *t*-tests (Bonferroni corrected) showed that all three ROIs showed a strong effect of *Periodicity* ($p < .0001$ in all cases), with higher responses in the periodic compared to nonperiodic conditions (see Fig. 3A). There were also significant main effects of both *Periodicity*, $F(2,34) = 95.19$, $p < .0001$, $\eta_G^2 = 0.66$, and *ROI*, $F(2,34) = 26.81$, $p < .0001$, $\eta_G^2 = 0.20$, however owing to the presence of a significant interaction, we did not interpret these main effects further.

4.2.2. Response at the common frequency (12 Hz)

Collapsing across all conditions and channels, we observed a large response at the exact frequency of stimulation (i.e., 12 Hz) and its harmonics 24 Hz and 36 Hz. The response at these frequencies peaked over medial occipital sites (around Oz, see Fig. 2C), a similar topography to that observed in previous studies using this stimulation rate (e.g., Retter and Rossion, 2016). Since this common response reflects neural synchronization to the stimulus onset/offset rate, it should not differ between the periodic and nonperiodic conditions. Indeed, the global scalp response at these three harmonics of 12 Hz was significant for both the periodic and nonperiodic conditions ($p < .001$, one tailed). Moreover, the sum of baseline-corrected amplitude values at these three frequencies did not differ between the periodic ($M = 0.713$ μV, $SD = 0.35$ μV) and nonperiodic conditions ($M = 0.685$ μV, $SD = 0.28$ μV), $t(17) = 0.519$, $p = 0.610$, $d = 0.122$. Taken together, these results provide no evidence that participants attended differently to the periodic or nonperiodic stimulation sequences.

Inspected as a function of ROI, all three regions showed a highly significant response ($p < .001$, one tailed) at the first three harmonics of the common frequency (i.e., 12, 24, and 36 Hz). Accordingly, we summed the response across these frequencies and subjected these values to a repeated measures ANOVA with the factors *Periodicity* and *ROI*. Here the interaction was not significant, $F(1.07,18.26) = 0.41$, $p = 0.544$, $\eta_G^2 = 0.0008$, nor was the main effect of *Periodicity*, $F(1,17) = 0.168$, $p = 0.687$, $\eta_G^2 = 0.0007$. This suggests that the magnitude of the common response was comparable across the periodic ($M =$

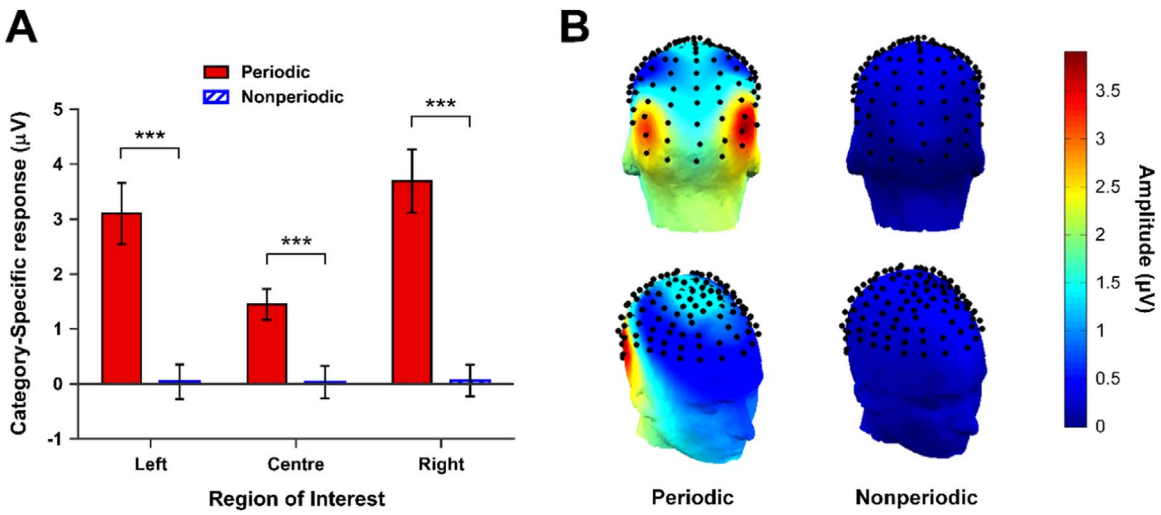


Fig. 3. The category-selective response in Expt. 1. **(A)** Sum of baseline-corrected amplitudes at the first 13 harmonics of 0.67 Hz, shown as a function of face periodicity and ROI. All three ROIs showed a strong category-selective response in the periodic condition, but not the nonperiodic condition. Error bars are within-subjects 95% CIs (overlap should not be interpreted by eye, see Cumming and Finch, 2005). Significance codes: *** $p < .001$ **(B)** Scalp topographies for the category-selective response in the periodic and nonperiodic conditions. Where the periodic condition was characterized by a strong bilateral occipito-temporal category-selective response, there was no significant category-selective response at any electrode site in the nonperiodic condition.

1.33 μV, $SD = 1.04$ μV) and nonperiodic ($M = 1.30$ μV, $SD = 0.93$ μV) conditions. In contrast, there was a highly significant effect of *ROI*, $F(1.39,23.70) = 36.39$, $p < .0001$, $\eta_G^2 = 0.48$. Posthoc Tukey HSD contrasts showed the common response was significantly larger in the central ROI ($M = 2.25$ μV, $SD = 1.08$ μV) compared to both the left ROI ($M = 0.66$ μV, $SD = 0.31$ μV), $z = -8.42$, $p < .001$, $d = 1.720$, and right ROI ($M = 1.04$ μV, $SD = 0.52$ μV), $z = -6.37$, $p < .001$, $d = 1.23$. In contrast, the response in the two lateral ROIs did not differ significantly, $z = 2.04$, $p = 0.102$, $d = 0.795$.

4.3. False-sequencing

Having verified that the periodicity manipulation was effective, we next turned to *quantifying* the response elicited by periodic and nonperiodic faces using false-sequencing (see Section 3 for details). As for the initial frequency-domain analysis, here we inspected the category-selective response (imposed in these false-sequences at exactly 1 Hz and harmonics) as well as the common response at 12 Hz and harmonics.

4.3.1. False-sequence response at the category-selective frequency (1 Hz)

Using scalp-averaged false-sequence data, we extracted *z*-scores for the imposed face frequency of 1 Hz and its harmonics under 12 Hz. In both the periodic and nonperiodic false-sequences, all harmonics in all ROIs met our significance criterion. To compare the magnitude of the scalp-averaged category-selective response across conditions, we summed the response across the first 11 harmonics of 1 Hz. A paired *t*-test showed there was no significant difference in response magnitude between the periodic ($M = 1.156$ μV, $SD = 0.515$ μV) and nonperiodic ($M = 1.104$ μV, $SD = 0.542$ μV) false-sequence conditions, $t(17) = 0.836$, $p = 0.415$, $d = 0.197$. Next, we extracted the baseline-subtracted amplitude spectra for the periodic and nonperiodic false-sequences in each ROI separately. Importantly, z-score inspection revealed a significant response at all 11 harmonics of 1 Hz in all three ROIs, regardless of whether the false-sequenced data came from the periodic or nonperiodic condition (Fig. 4).

To quantify the response to periodic and nonperiodic faces, we summed the baseline-corrected amplitude values at the 11 harmonics for each condition/ROI pair (Fig. 5A), and subjected these data to a repeated measures ANOVA with the factors *Periodicity* and *ROI*. Here the interaction between these factors did not approach significance, $F(2,34) = 1.0$, $p = 0.379$, $\eta_G^2 = 0.0007$, nor did the main effect of *Periodicity*, $F(1,17) = 2.79$, $p = 0.113$, $\eta_G^2 = 0.004$. As such, there was no evidence that category-selective response differed between the periodic ($M = 2.76$ μV; $SD = 1.55$ μV) and nonperiodic false-sequences ($M = 2.58$ μV; $SD = 1.61$ μV). In contrast, there was a significant main effect of *ROI*, $F(2,34) = 22.87$, $p < .0001$, $\eta_G^2 = 0.24$. Posthoc Tukey HSD contrasts indicated the response in the right ROI ($M = 3.51$ μV, $SD = 1.57$ μV) was greater than the response in both the center ROI ($M = 1.65$ μV; $SD = 0.835$ μV), $z = 6.86$, $p < .001$, $d = 1.508$, and left ROI ($M = 2.85$ μV; $SD = 1.61$ μV), $z = 2.43$, $p = 0.04$, $d = 0.612$. Also, the response in the left ROI was greater than that in the center ROI, $z = 4.43$, $p < .001$, $d = 0.975$.

Since we were conscious of the problems associated with drawing inferences from *p*-values in support of the null hypothesis (Dienes, 2014), we also calculated a Bayes factor for each model combination involving the factors *Periodicity* and *ROI*. This analysis showed that the *ROI*-only model provided the best fit, accounting for our data approximately 29,620 times better than the null model (i.e., intercept only). In contrast, the Bayes factor for the *Periodicity*-only model was just 0.238, suggesting the null model accounted substantially better for our data than one including just *Periodicity* (Jeffreys, 1939/1961). Furthermore, when comparing our preferred *ROI*-only model with all other model combinations, it performed approximately 4.05 times better than the next best model which included additive effects of both *ROI* and *Periodicity*. Thus, the inclusion of *Periodicity* did not add sufficient explanatory power to overcome Bayesian penalties for increasing model complexity.

Given the similarity in the magnitude of the response to periodic and nonperiodic faces, we expected there would be a strong correlation between the category-selective response magnitudes in the two conditions. Fig. 6 shows this was indeed the case – there was a very high correlation between the face response for the periodic and nonperiodic conditions in both the left ROI, $r(16) = 0.931$, $p < .0001$, and the right ROI, $r(16) = 0.935$, $p < .0001$. In short, individuals with a strong face-selective response in the periodic condition also had a strong face-selective response in the nonperiodic condition. Similarly strong correlations were observed for participants' response significance values (i.e., *z*-scores) in both the left ROI, $r(16) = 0.829$, $p < .0001$, and the right ROI, $r(16) = 0.710$, $p = 0.0010$.

4.3.2. False-sequence response at the common frequency (12 Hz)

Where the false-sequencing procedure imposes the category-

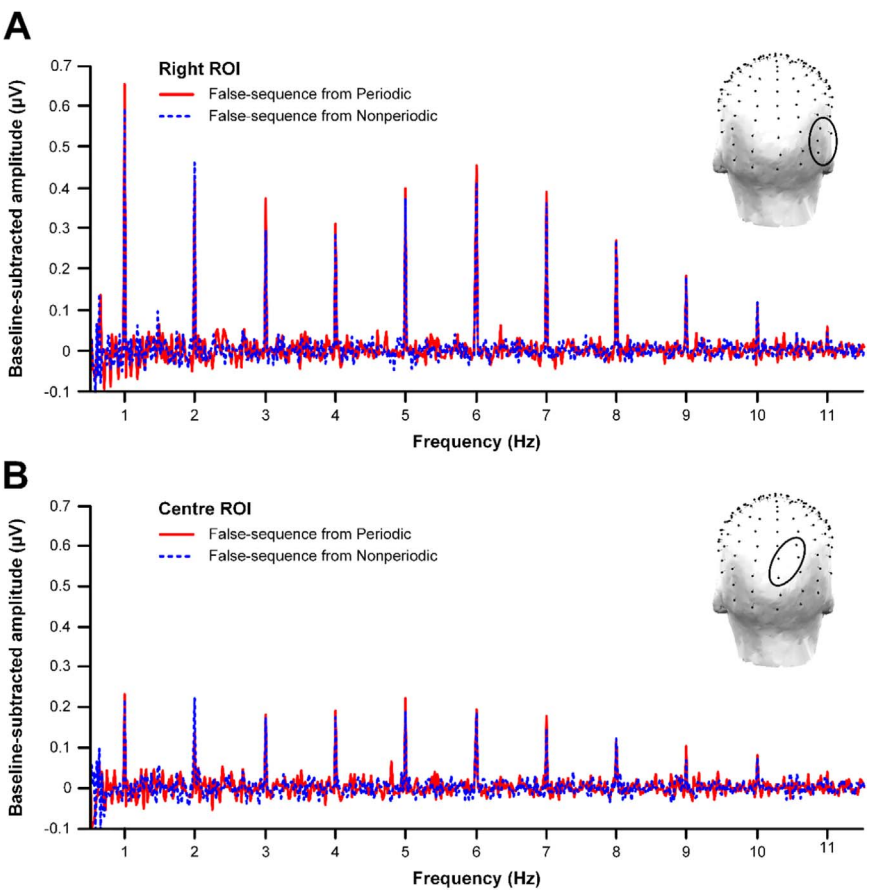


Fig. 4. Baseline-subtracted amplitude spectra for the false-sequences corresponding to the periodic and nonperiodic conditions, shown here for the **(A)** right ROI and **(B)** center ROI. Note that unlike the original frequency-domain analyses, the false-sequences generated from both periodic and nonperiodic sequences showed a significant response at the category-selective frequency (i.e., 1 Hz) and its harmonics.

selective response at a specified frequency (i.e., 1 Hz), it should have no impact on the frequency at which the common response is observed (i.e., 12 Hz). Indeed, collapsing across all channels and conditions, our *z*-score extraction procedure identified a highly significant response for both false-sequence types at 12 Hz, 24 Hz, and 36 Hz ($p < .001$, one tailed). Accordingly, we subjected the sum of baseline-corrected amplitudes at these three frequencies for each condition/ROI pair to a repeated measures ANOVA with the factors *Periodicity* and *ROI*. Here we observed an identical pattern of results to that seen in the original frequency-domain analysis (see Fig. 5B). That is, the *Periodicity* × *ROI* interaction was not significant, $F(2,34) = 1.83$, $p = 0.176$, $\eta_G^2 = 0.0002$, nor was the main effect of *Periodicity*, $F(1,17) = 0.877$, $p = 0.362$, $\eta_G^2 = 0.0001$. Thus, there was no evidence that the magnitude of the common response differed for periodic ($M = 1.44$ μV; $SD = 1.10$ μV) and nonperiodic ($M = 1.42$ μV; $SD = 1.09$ μV) false-sequences. In contrast, there was a strong main effect of *ROI*, $F(1.27,21.62) = 32.63$, $p < .0001$. Posthoc Tukey HSD contrasts indicated this was due to significantly lower responses in the left *ROI* ($M = 0.720$ μV; $SD = 0.303$ μV) and right *ROI* ($M = 1.10$ μV; $SD = 0.519$ μV) compared to the center ROI ($M = 2.48$ μV; $SD = 1.23$ μV), Center-Left: $z = -7.90$, $p < .0001$, $d = 1.565$; Center-Right: $z = -6.19$, $p < .0001$, $d = 1.191$. In contrast, the common response in the right and left ROIs was not significantly different, $z = 1.71$, $p = 0.203$, $d = 0.790$.

4.4. Time-domain results

We further examined the response elicited by periodic and nonperiodic faces in the time-domain. Fig. 7 shows the mean waveforms for all 128 channels after selectively removing the image presentation frequency of 12 Hz (see Fig. S1 in Supplemental materials for time-domain responses in which the 12 Hz frequency is preserved). This response was remarkably similar in the two periodicity conditions, notably characterized by the same four distinct spatiotemporal components we have reported previously for a different set of faces in an FPVS-EEG design (Retter and Rossion, 2016).

We compared the response to temporally predictable and unpredictable faces in each ROI by conducting a paired *t*-test at each of the 513 samples. Any observed *t*-value that exceeded the critical cut-offs given by our conservative permutation procedure (see Section 3) was considered a significant difference between the periodic and nonperiodic conditions. As can be seen in Fig. 8, no observed *t*-value in any ROI approached either this permutation defined cut-off, or even a less conservative cut-off of $p < .01$, uncorrected. As such, there was no evidence to suggest that the response evoked by periodic and nonperiodic faces in the fast periodic visual stream differed meaningfully in any ROI.

To further assess any difference in the response evoked by periodic and nonperiodic faces, we also subjected the time-domain data to MVPA decoding. In the first of these analyses, we trained an SVM classifier on the spatiotemporal activity patterns associated with periodic and nonperiodic faces in 20 ms increments (−166 ms to 500 ms, see Section 3 for details). Paired *t*-tests ($n = 34$) indicated this classifier was not able to identify novel, unlabeled face responses as either periodic or nonperiodic significantly better than empirical chance at any time point following face onset. A complementary question was whether a classifier trained to distinguish objects and faces in the periodic condition would generalize well to the same distinction in the nonperiodic condition. To this end, we trained a second SVM classifier on spatiotemporal patterns of activation elicited by face and object images in the periodic condition (−166 ms to 500 ms, see Section 3 for details). We then tested classifier performance using novel data drawn from either the periodic condition (i.e., within-condition decoding), or the nonperiodic condition (i.e., cross-condition decoding). Here

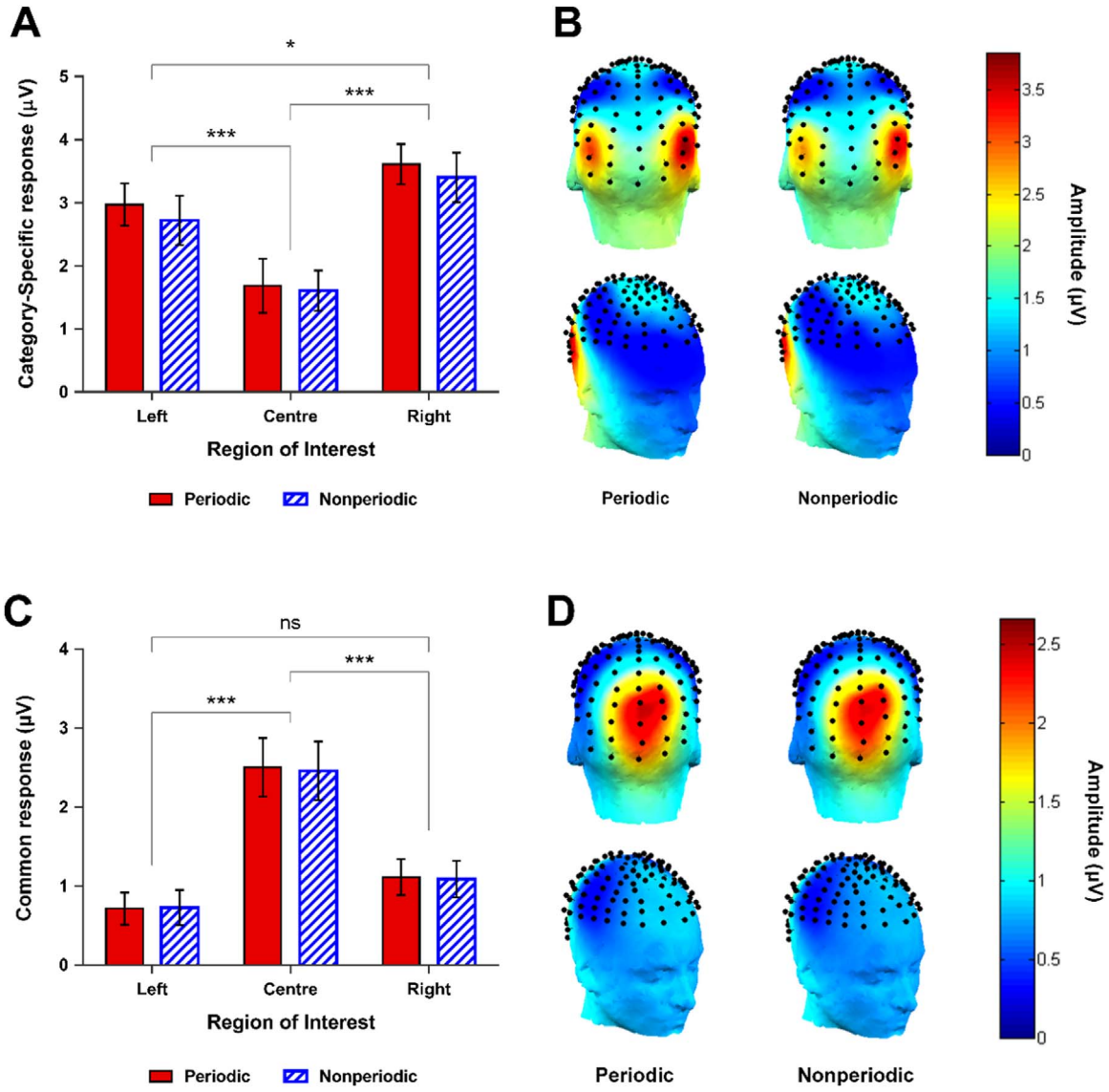


Fig. 5. Sum of baseline-corrected amplitudes representing the **(A)** category-selective response (first 11 harmonics of 1 Hz) and **(C)** common response (first 3 harmonics of 12 Hz) for the false-sequence data in Expt. 1. Both the category-selective and common response were strongly modulated by ROI, but were not at all sensitive to face periodicity. Error bars are within-subjects 95% CIs (overlap should not be interpreted by eye, cf. Cumming, 2005). Significance codes: *** $p < .001$; ** $p < .01$; * $p < .05$. **(B)** Scalp topographies corresponding to the category-selective response and **(D)** common response, shown as a function of face-periodicity.

classifier decoding accuracy for the within-condition analysis acts as a baseline, reflecting the degree to which the response to faces can be reliably distinguished from the response to objects. Indeed, within-condition decoding of face and object responses was significantly higher than empirical chance from ~ 137 ms to ~ 352 ms following stimulus onset, with peak decoding at ~ 274 ms post onset (Fig. 9A, blue lines). Importantly, cross-condition decoding accuracy was also significantly above empirical chance from ~ 118 ms to ~ 371 ms after stimulus onset (Fig. 9A, red lines), with peak decoding at exactly the same time point (~ 274 ms) as in the within-condition decoding analysis. Moreover, two-tailed paired *t*-tests (Bonferroni corrected) indicated decoding accuracy was not significantly different in the within-condition and cross-condition analyses at any time point, suggesting the response evoked by faces in the periodic condition generalized equally well to novel nonperiodic data as it did to novel periodic data. Indeed if anything, cross-condition decoding started slightly earlier and lasted longer than within-condition decoding, even though there was no significant difference in decoding accuracy at any time point.

5. Experiment 1 interim discussion

In Expt. 1 we set out to compare the category-selective response evoked by faces embedded at temporally predictable or unpredictable intervals in an FPVS sequence. We found no evidence of a qualitative or quantitative difference in the response to temporally predictable and unpredictable faces. In both conditions, the category-selective response was comprised of the same four spatiotemporal components we have reported previously for another set of natural face images in FPVS sequences (Jacques et al., 2016; Retter and Rossion, 2016): P1-face (peaking at 145 ms), N1-face (peaking at 220 ms), P2-face (peaking at 290 ms), and P3-face (peaking at 465 ms). Several results suggest that this ~ 400 ms category-selective response is immune to the temporal predictability of faces. First, our false-sequencing quantification showed that the *magnitude* of the category-selective response was not sensitive to face periodicity, although it was modulated by ROI (right > left > center). Moreover, a Bayes factor analysis of response magnitudes indicated that the null model was strongly preferred over one including face periodicity. Second, permutation test procedures across the 800 ms

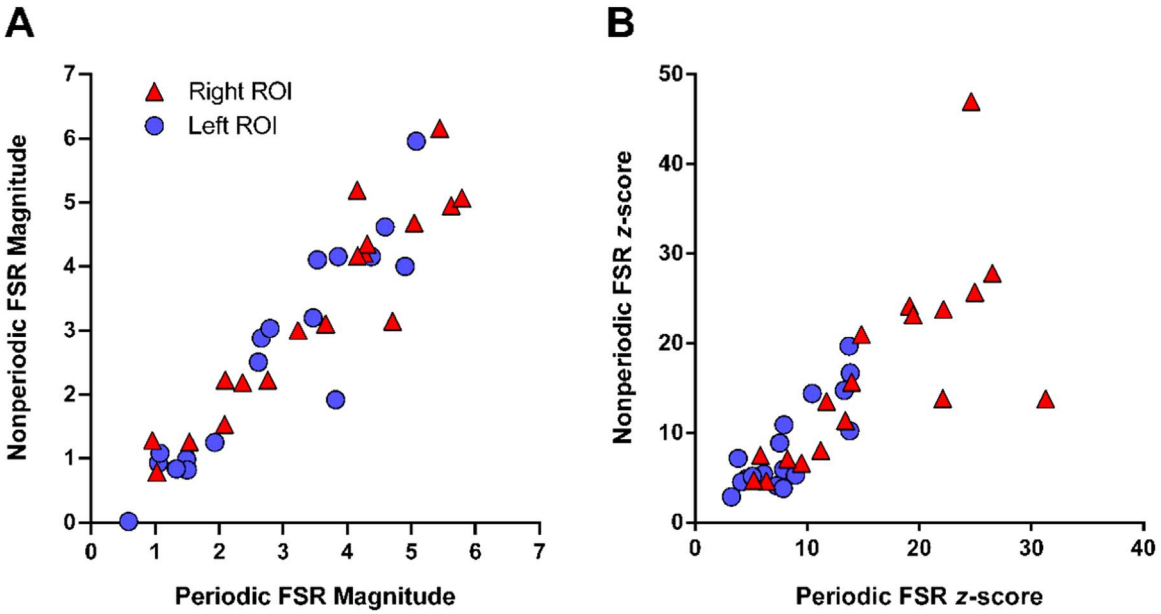


Fig. 6. Correlations in category-selective response properties for the periodic and nonperiodic false-sequences **A)** Response magnitude correlations for the two lateral ROIs (i.e., sum of baseline-corrected amplitudes across first 11 harmonics of 1 Hz). **B)** Response significance correlations (i.e., z-scores) for the two lateral ROIs.

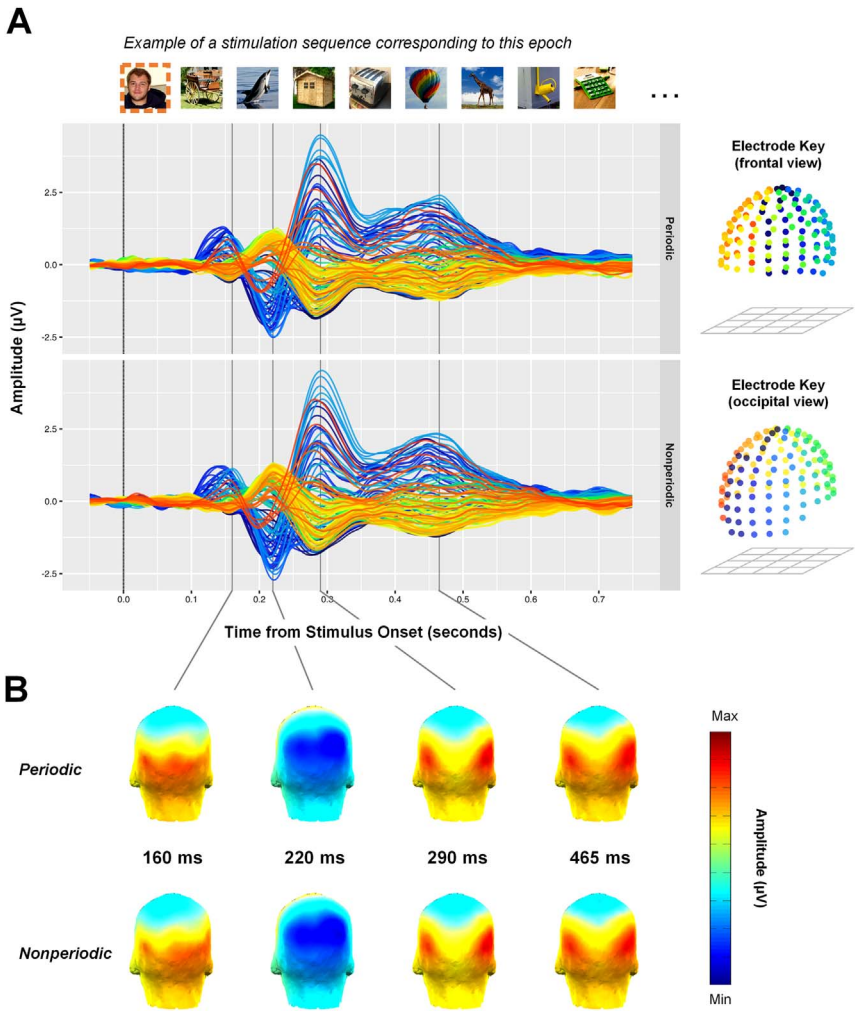


Fig. 7. Time-domain representation of the category-selective response elicited by faces in the periodic and nonperiodic conditions. **(A)** Conditional mean waveforms for all 128 channels plotted as a function of time from face onset. The image presentation frequency (12 Hz) has been selectively removed from these data using an FFT multi-notch filter (see Fig. S1 in Supplemental materials for time-domain responses preserving the 12 Hz response, also Fig. 4 in Retter and Rossion, 2016). **(B)** Scalp topographies for the periodic and nonperiodic conditions corresponding to the four spatiotemporal peaks comprising the category-selective response (0.165 s, 0.220 s, 0.290 s, and 0.465 s). Amplitude scales are fixed for each condition pair, but vary across time points.

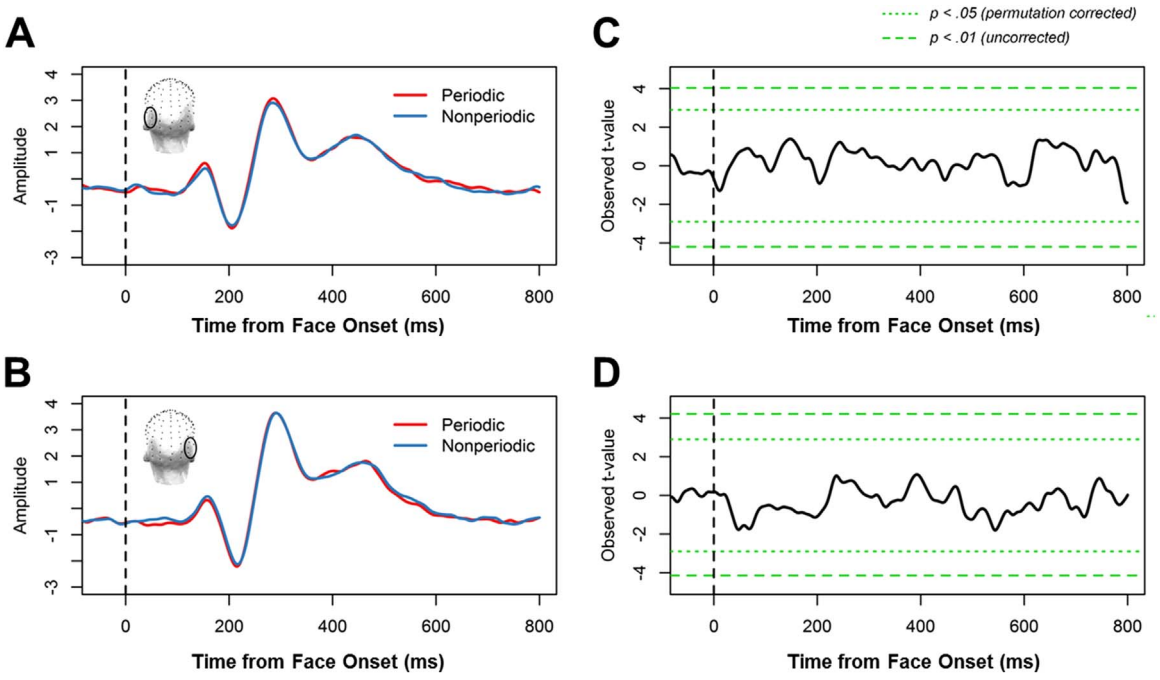


Fig. 8. Time-domain analysis results for Expt. 1. Left column: Conditional mean amplitude after selectively removing the response at 12 Hz, shown as a function of time from face onset for the **(A)** left ROI and **(B)** right ROI. Right column: The observed *t*-value for each time point, shown for the **(C)** left ROI and **(D)** right ROI. Horizontal dashed lines reflect criterion significance cut-offs (permutation corrected and uncorrected). No observed *t*-value at any sample in any ROI approached either criterion cut-off value, suggesting there was no difference in the response evoked by periodic and nonperiodic faces.

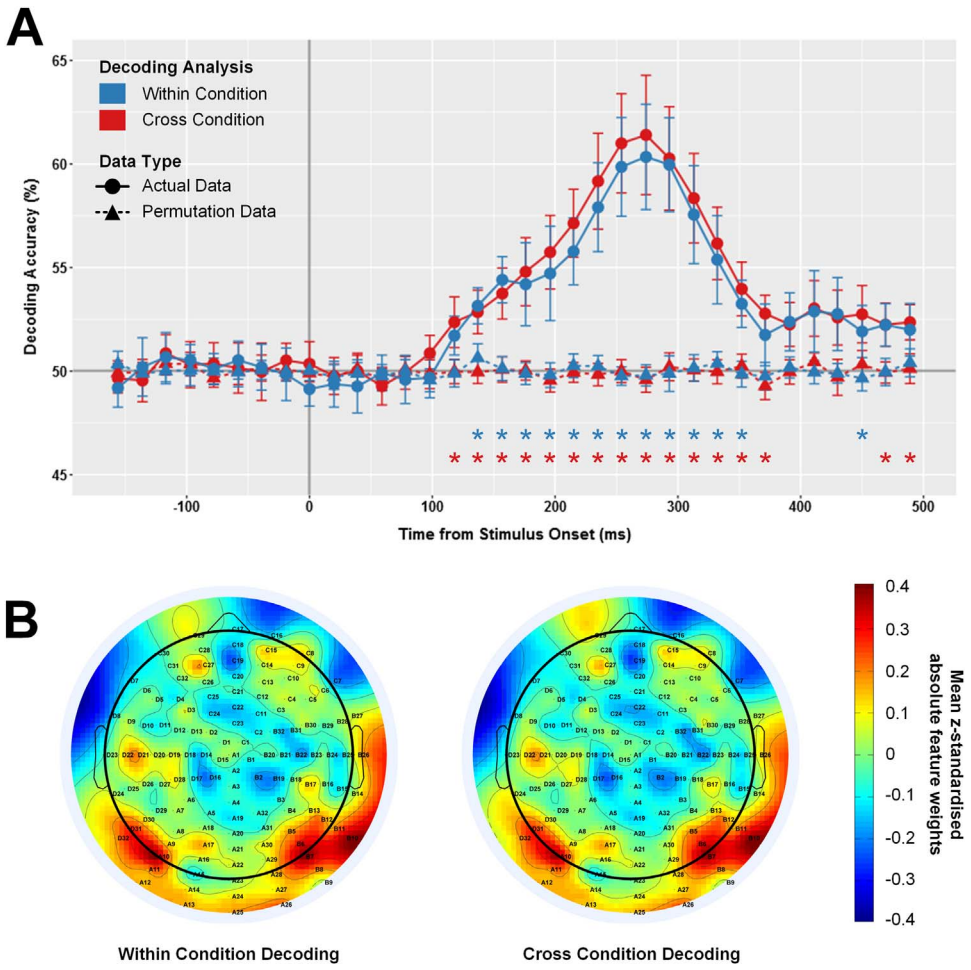


Fig. 9. (A) Face vs. object decoding accuracies for the within-condition (blue) and the cross-condition (red) analyses, shown as a function of data type (actual and permutated). Error bars represent SEM; asterisks indicate significantly higher decoding accuracy for the actual data compared to the permutated control data ($p < .05$, Bonferroni corrected). Note that decoding accuracy was not significantly different for the within-condition and cross-condition analyses at any time point **(B)** *z*-standardized absolute feature weights for all 128 channels, averaged across 150–350 ms, shown for the within-condition and cross-condition analyses. In both cases, the primary sources of face vs. object information were bilateral occipito-temporal channels. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article).

following face onset found no difference in the time-domain response to periodic and nonperiodic faces in any of our ROIs, nor was a linear classifier trained on periodic and nonperiodic face responses able to classify unlabeled responses better than empirical chance. Importantly, however, a different classifier trained to differentiate the response evoked by faces and objects in the periodic condition was able to classify unlabeled face/object responses equally well regardless of whether they were drawn from the periodic or nonperiodic condition. This latter result suggests that the category-selective response was qualitatively similar in the two periodicity conditions. Taken together, these results suggest that under orthogonal task conditions (as used in typical studies with this paradigm), the category-selective response in FPVS designs is not generated, or even modulated, by the temporal predictability of critical category exemplars.

Why did periodic and nonperiodic faces in our study elicit such similar responses when several existing studies have found that rhythmic temporal stimulation does indeed modulate perceptual processing (Cravo et al., 2013; Breska and Deouell, 2014; Morillon et al., 2016)? One possibility is that temporal expectation effects were masked here by an overall attentional imbalance between the conditions. Specifically, if observers paid more attention to the FPVS sequences in one condition compared to the other, then the category-selective signal in the better-attended condition could have been boosted so that it equaled that in the other condition, effectively concealing an effect of temporal predictability. If that were the case, however, there should be a commensurate boost to the response at the common frequency, as "steady-state visual evoked potential" (SSVEP) responses at such high frequency rates are known to substantially increase with focused selective attention (e.g., Müller et al., 2006). We found no such conditional difference in the 12 Hz response magnitudes, suggesting participants attended to sequences equally well regardless of face periodicity. As such, a difference in attentional allocation seems unlikely to account for why temporal predictability did not modulate the category-selective response in our study.

An alternative possibility is that, owing to the high saliency of faces for the human brain (Hershler and Hochstein, 2005; Crouzet et al., 2010), participants' temporal expectations about the predictable onset of faces were not sufficient to modulate the strong face-selective evoked response. That the neural response to faces might be robust to modulation by temporal expectation is not an unreasonable possibility, considering that effects of factors such as spatial attention on face-processing have historically been quite difficult to find (Reddy et al., 2004, 2006; Finkbeiner and Palermo, 2009; Quek and Finkbeiner, 2013). If is indeed the case, then another way to examine whether temporal predictability drives the category-selective response in FPVS is to observe the neural response to a violation of temporal expectations, i.e., when the visual stimulus encountered does not match a pre-activated template (Rao and Ballard, 1999; Friston, 2005; Kok et al., 2014). We tested this hypothesis in Experiment 2.

6. Experiment 2

To examine the neural response to a violation of rhythmic temporal expectations, participants in Expt. 2 viewed rapid continuous sequences of natural object images with a face embedded as every 12th image (i.e., at periodic intervals in the sequence). After an initial period of familiarization, a small fraction (10%) of these highly temporally predictable faces were replaced with a randomly selected object, a stimulus we refer to here as a "missing face". If temporal expectations about faces contribute to the category-selective signal in FPVS, then non-face stimuli that violate these expectations (i.e., missing faces) should elicit a different response than non-face stimuli about which participants cannot build any temporal expectations. In simple terms, the response evoked by an object that appears in place of an expected face should differ in some way from that evoked by the other objects in the sequence. If this is indeed the case, a false-sequencing analysis of missing faces will capture this differential response.

Importantly, our approach here makes no assumptions about the nature of the differential response to missing faces. It could be, for example, that objects replacing expected faces evoke a "prediction error" response, perhaps similar in kind to the mismatch negativity (MMN; for a review, see Naatanen et al., 2010) that has been associated in recent theoretical accounts with an error detection signal (Friston, 2005; Garrido et al., 2009; Kimura et al., 2011; Stefanics et al., 2011; Lieder et al., 2013; Pieszek et al., 2013). On the other hand, perhaps participants' temporal expectations will drive a partly face-like selective response to missing faces, even in the absence of a true face stimulus. Indeed previous studies have shown that unexpected omissions of visual stimuli can elicit feature-specific responses in visual cortex (den Ouden et al., 2009; Kok et al., 2014). Whatever its exact nature, if the response evoked by objects replacing expected faces differs in any way from that evoked by other objects in the sequence, it will be reflected in a significant signal at the imposed frequency in a false-sequence.

7. Experiment 2 methods

7.1. Participants

A different group of 13 individuals (eight females) aged between 19 and 26 years took part in Expt. 2 in exchange for payment. Participation criteria were the same as Expt. 1; we obtained written informed consent from all participants prior to testing.

7.2. Procedure

The stimuli, experimental testing set up, and fixation cross task were identical to those described in Expt. 1. Each trial consisted of a pre-stimulation period containing only the central fixation cross (2 s), a fade-in (2 s), the stimulation sequence itself (120 s), and a fade-out (3 s). As in Expt. 1, the stimulation sequence was a series of randomly selected object images presented at a periodic rate of 12 Hz. Here we embedded a natural face image after every 11 object images, yielding a face periodicity of exactly 1 Hz (i.e., 12 Hz image presentation rate/12 images). The critical manipulation was the inclusion of "missing faces" within each sequence, achieved by replacing 10% of the periodic faces with a randomly selected object image (Fig. 10).[1] The first missing face could occur only after participants had viewed at least 10 normal face presentation cycles (i.e., 10 s of familiarization), after which missing faces were separated by a variable delay of 6–15 normal face presentation cycles (i.e., 6–15 s). There were 10 sequences, each containing approximately 108 real faces and up to 12 missing face instances, so that the conditions were optimal for temporal expectation.

7.3. EEG analysis

7.3.1. EEG pre-processing

We excluded data from two female participants whose EEG trace on initial inspection contained excessive artefact that unduly affected time-domain visualization. As in the case of the participants removed in Expt. 1, both individuals nevertheless showed a significant response at the face periodicity and several harmonics (1 Hz). All remaining pre-processing steps were identical to those implemented in Expt. 1, save that here we segmented longer epochs to correspond to the 120 s sequence.

7.3.2. Frequency-domain analysis

The goal of the initial frequency analysis was to verify that FPVS

[1] Note that unlike some previous studies (den Ouden et al., 2009; Kok et al., 2014) in which a stimulus omission means there is no visual input at all, here the expected stimulus (a face) is *replaced* with an unexpected stimulus (an object).

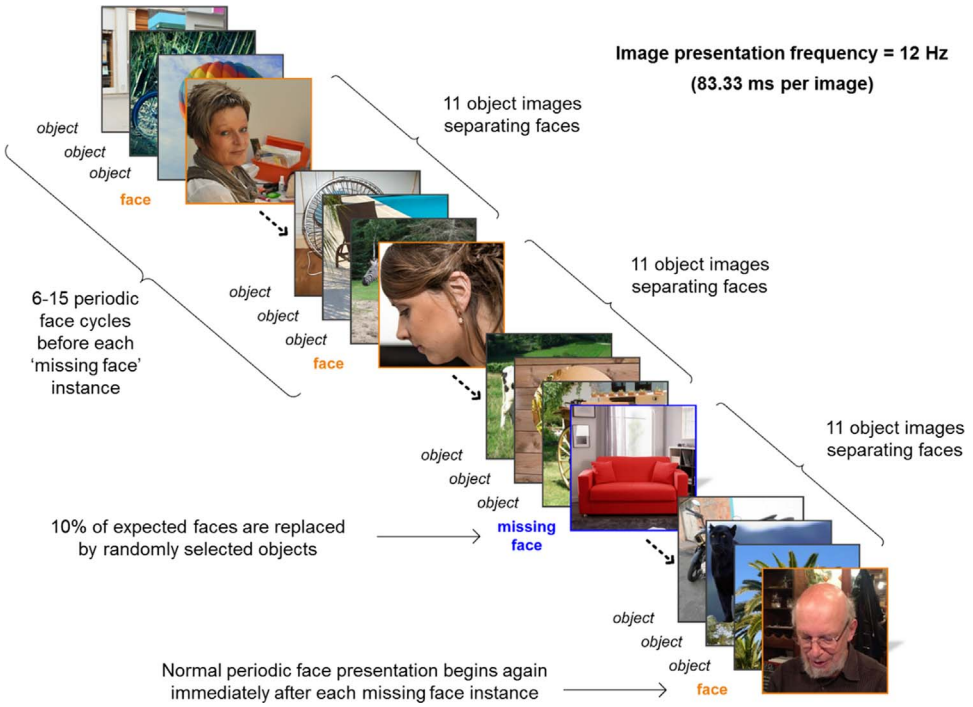


Fig. 10. Schematic representation of the stimulation sequence for Expt. 2. Natural object images appeared at a periodic rate of 12 Hz, with a face embedded as every 12th image (i.e., face periodicity = 1 Hz). Within each sequence, 10% of periodic faces were replaced by a randomly selected object image, resulting in a "missing face" for the observer. 6–15 normal periodic face cycles could occur between missing face instances. As in Expt. 1, the task was to monitor a central fixation cross which overlaid the images for changes of color.

sequences in which 10% of periodic faces were removed nevertheless still evoked a category-selective response (i.e., response at 1 Hz). To this end, we cropped the EEG recording for each sequence to a 119.02 s epoch containing exactly 119 face presentation cycles, and subjected each participants average epoch to a Fast Fourier Transformation (FFT). The frequency resolution of the resulting normalized amplitude spectra was 0.0084 Hz (the inverse of the sequence duration, i.e., 1/119.02). We identified significant responses at the frequencies of interest by pooling the amplitude spectra across channels and computing a *z*-score at each frequency bin (significance criterion = $z > 3.1$, or $p < .001$, one-tailed, i.e., signal > noise), using the same parameters as applied for Expt. 1. We then quantified the response at the relevant frequencies in three ROIs defined using the same procedures as for Expt. 1 (i.e., by summing baseline-corrected amplitudes). This procedure yielded left and right occipito-temporal ROIs that were identical to those of Expt. 1 (left: P9, PO7, PO9, & PO11; right: P10, PO8, PO10, & PO12), and a near-identical central ROI (O2, POOz, PO4h, POO6).[2]

7.3.3. False-sequencing analysis

Our goal here was to quantify and compare the response to real and "missing" faces using the same false-sequencing procedure described for Expt. 1. Epochs were 1000 ms long, lasting from 500 ms before to 500 ms after each missing face instance, resulting in ~ 12 segments that we sequenced, aligned, linearly de-trended, and cropped to exactly 11 s. We repeated this same procedure to produce false-sequences corresponding to each real face occurrence immediately preceding a missing face. This ensured an equal number of real and missing faces in the false-sequencing analysis, in which the imposed critical stimulus periodicity was exactly 1 Hz. We averaged the resulting false-sequences by participants before applying an FFT. The baseline correction was in Expt. 1, save that owing to the lower frequency resolution here (i.e., 1/11 = 0.0113 Hz), the noise range comprised 18 neighboring frequency bins (excluding the immediately adjacent bins and the local maximum and minimum amplitude bins).

[2] For completeness' sake we also inspected the common response magnitude in a central ROI identical to that used in Expt. 1 (i.e., averaged over POOz, Oz, POO6, PO4h), and found it to be just 0.0055 μV lower than that given by the central ROI we used in the main analyses (i.e., averaged over POOz, O2, POO6, PO4h).

7.3.4. Time-domain analysis

We inspected the EEG waveform data time-locked to a comparable number of missing faces and real faces. Here we applied a zero phase-shift Butterworth low-pass filter (4th order, 30 Hz cut-off) to the continuous EEG trace, before cropping each sequence to an integer number of 12 Hz cycles. We applied an FFT multi-notch filter (Hanning window, width = 0.05 Hz) to selectively remove the response at the common frequency of 12 Hz, and three of its harmonics (24 Hz, 36 Hz, and 48 Hz). We then segmented an epoch lasting from −200 ms to 750 ms around *i)* each missing face in the sequence, and *ii)* each real face instance immediately preceding a missing face. We baseline-corrected each epoch to the average of the 166 ms preceding stimulus onset, and averaged the resulting ~ 120 epochs of each type to create conditional means. A two-tailed paired *t*-test between the condition of interest and a baseline of zero at each time point ($n = 564$) was computed to assess the presence of a category-selective response for each condition. As in Expt. 1, we inspected the obtained *t*-values relative to two significance criteria ($p < .01$ uncorrected, and $p < .05$, corrected for multiple comparisons). We obtained the second, more conservative criterion using the same permutation procedure described for Expt. 1 (2048 permutations implemented across 564 time points, for a total of 1.16 million comparisons).

8. Experiment 2 results

8.1. Behavior

Detection accuracy for the fixation cross color change task was near ceiling at 95.85% (SD = 6.02%). The mean response time (RT) on correct trials was 502 ms (SD = 42 ms).

8.2. Frequency-domain results

8.2.1. Response at the category-selective frequency (1 Hz)

Z-score inspection at the scalp-average level revealed a significant response at each of the first 18 harmonics of 1 Hz. Thus, despite the inclusion of missing faces, the stimulation sequences nevertheless still evoked a strong category-selective response. A repeated measures ANOVA indicated that the sum of these 18 harmonics (excluding the

12th harmonic, i.e., the 12 Hz common frequency) varied significantly with *ROI*, $F(2,20) = 10.86$, $p < .001$, $\eta_G^2 = 0.19$. Posthoc Tukey HSD contrasts indicated that the right ($M = 3.45$; $SD = 1.76$) and left ROIs ($M = 2.74$; $SD = 1.71$) exhibited a significantly larger category-selective response than the central ROI ($M = 1.76$; $SD = 0.904$) (right vs. center: $z = 4.87$, $p < .001$, $d = 1.164$; left vs. center: $z = 2.82$, $p = 0.013$, $d = 0.756$). However, these two lateral regions did not differ significantly from one another, $z = 2.04$, $p = 0.102$, $d = 0.927$.

8.2.2. Response at the common frequency (12 Hz)

Again collapsing across all channels, there was also a highly significant response at the first five harmonics of the common frequency (i.e., 12, 24, 36, 48, & 60 Hz). We examined the sum of baseline-subtracted amplitude values across these five harmonics in each of predefined ROI. A repeated measures ANOVA here revealed a highly significant effect of *ROI* on the magnitude of the common response, $F(1.16, 11.62) = 25.02$, $p < .001$, $\eta_G^2 = 0.55$. As was the case in Expt. 1, posthoc Tukey HSD contrasts showed this response was significantly larger in the central ROI ($M = 2.13\,\mu V$, $SD = 0.978\,\mu V$) compared to both the left ROI ($M = 0.562\,\mu V$, $SD = 0.246\,\mu V$), $z = -7.09$, $p < .0001$, $d = 1.522$, and right ROI ($M = 0.926\,\mu V$, $SD = 0.426\,\mu V$), $z = -5.44$, $p < .0001$, $d = 1.679$. In contrast, there was no significant difference in the magnitude of the common response between the left and right ROIs, $z = 1.65$, $p = 0.224$, $d = 0.814$.

8.3. False-sequencing

The false-sequencing procedure generated 12 Hz sequences containing either a "missing face" or real face at an imposed periodicity of 1 Hz. We inspected the response at both the category-selective frequency and common frequency for these two sequence types.

8.3.1. False-sequence response at the category-selective frequency (1 Hz)

Pooling across all scalp channels and conditions, we identified significant responses ($p < .001$, one tailed) at the 4th–8th harmonic of the imposed periodicity at 1 Hz. As such, we summed the baseline-corrected amplitude values for each condition across the initial eight harmonics of 1 Hz.[3] Averaged across the whole scalp, this aggregated category-selective response was significantly larger for the real face sequences ($M = 0.526\,\mu V$, $SD = 0.391\,\mu V$) compared to the missing face sequences ($M = -0.171\,\mu V$, $SD = 0.150\,\mu V$), $t(10) = 5.60$, $p < .001$, $d = 1.688$).

Taking the aggregated category-selective response in each ROI, a repeated measures ANOVA with the factors *Sequence Type* and *ROI* revealed a significant interaction, $F(2,20) = 4.59$, $p = 0.023$, $\eta_G^2 = 0.07$, the nature of which is clear in Fig. 11A. As in Expt. 1, *ROI* significantly modulated the category-selective response in the real face condition, but not in the missing face condition (since responses in the latter were at floor level). up pairwise *t*-tests indicated there was a significant effect of *Sequence Type* in each ROI (all Bonferroni adjusted *p* values < 0.05). There was also a significant main effect of *Sequence Type*, $F(1,10) = 33.42$, $p = < 0.001$, $\eta_G^2 = 0.48$, and *ROI*, $F(2,20) = 5.54$, $p = < 0.05$, $\eta_G^2 = 0.10$.

Since we were primarily interested in asking whether the missing faces generated any category-selective response *at all*, we also inspected the 95% confidence interval (CI) around the mean category-selective response for missing faces, and noted that it included zero (−0.493 to 0.148) where the 95% CI for real faces did not 1.170–1.812). This would suggest that real faces gave rise to a category-selective (i.e., differential) response ($M = 1.49\,\mu V$; $SD = 1.29\,\mu V$), where missing faces did not ($M = -0.172\,\mu V$; $SD = 0.389\,\mu V$) – a pattern that held for each participant when considered at the individual level (see Fig. S2 in Supplemental materials).

[3] Although the first three harmonics of 1 Hz did not reach our significance criterion, we had no a priori reason to exclude the signal on these harmonics from our quantification.

8.3.2. False-sequence response at the common frequency (12 Hz)

Again pooling across all channels and conditions, z-score inspection revealed a significant response at each of the first four harmonics of the common frequency (i.e., 12 Hz). At the level of ROI, a repeated measures ANOVA with the factors *Sequence Type* and *ROI* revealed that there was no interaction between these factors, $F(2,20) = 0.876$, $p = 0.432$, $\eta_G^2 = 0.001$, nor was there an effect of *Sequence Type*, $F(1,10) = 2.98$, $p = 0.115$, $\eta_G^2 = 0.003$. As can be seen in Fig. 11C, the magnitude of the common response was comparable for missing face sequences ($M = 1.16\,\mu V$; $SD = 0.918\,\mu V$) and real face sequences ($M = 1.10\,\mu V$; $SD = 0.926\,\mu V$). In contrast, there was a strong main effect of *ROI*, $F(1.16,11.62) = 23.65$, $p < .001$, $\eta_G^2 = 0.52$. Posthoc Tukey HSD contrasts showed that the common response for the central ROI ($M = 2.03\,\mu V$; $SD = 0.973\,\mu V$) was significantly larger than both the left ROI ($M = 0.479\,\mu V$; $SD = 0.270\,\mu V$), $z = -6.95$, p < .0001, $d = 1.509$, and right ROI ($M = 0.884\,\mu V$; $SD = 0.464\,\mu V$), $z = -5.14$, $p < .0001$, $d = 1.682$. However, there was no difference between the two lateral ROIs, $z = 1.82$, $p = 0.164$, $d = 0.817$.

8.4. Time-domain results

Fig. 12 shows the global response in the time-domain to missing and real faces, both before and after selectively removing the common response at 12 Hz. Where the category-selective response for real faces was comprised of the same four spatiotemporal components as in Expt. 1 (Retter and Rossion, 2016), the category-selective response to missing faces was minimal. That is, the response evoked by missing faces appeared to be almost entirely captured by the common response at 12 Hz.

For each condition separately, we assessed the presence of a category-selective response in each ROI by comparing the response with a zero baseline in a series of two-tailed paired *t*-tests. As can be seen in Fig. 13, where the components evoked by real faces met our imposed significance criteria (see Section 7), the response to missing faces did not approach either cut-off in any ROI. However, while the averaged response of the four channels in the right ROI was not significantly different from a zero baseline, an inspection of Fig. 12A suggests there might be a small increase on several occipito-temporal channels in the missing face condition around 300 ms. Given that this "blip" temporally coincides with the P3-face component, for interests' sake we further inspected the response on individual channels. One channel, PO12, met our less conservative significance criterion (i.e., $p < .01$, uncorrected) for a brief duration of 13.67 ms around 277 ms. Thus, on balance these time-domain analyses suggest that missing faces did not evoke an identifiable differential response. In other words, there was no evidence that objects replacing expected faces were processed differently to any other object in the rapid sequence.

9. Experiment 2 interim discussion

In Expt. 2 we examined whether rare omissions of highly temporally predictable faces in an FPVS sequence would evoke an expectation related response. If so, then the response to object images that replace expected faces (so-called "missing faces") should differ in some way from the response evoked by other object images in the sequence. The results here do not support this suggestion. Frequency-domain quantification of the differential response showed that where real faces elicited a strong category-selective response, the analogous response to missing faces was at floor level, i.e., not significantly greater than zero. Moreover, time-domain analysis showed that the evoked response to missing faces was entirely captured by the 12 Hz common response. This suggests that the response to objects replacing expected faces was no different to that of any other object in the sequence. Based on these results, we conclude that temporal expectation induced by the periodicity of critical stimuli in FPVS sequences is not sufficient to generate the category-selective response, at least when participants are engaged in an orthogonal task.

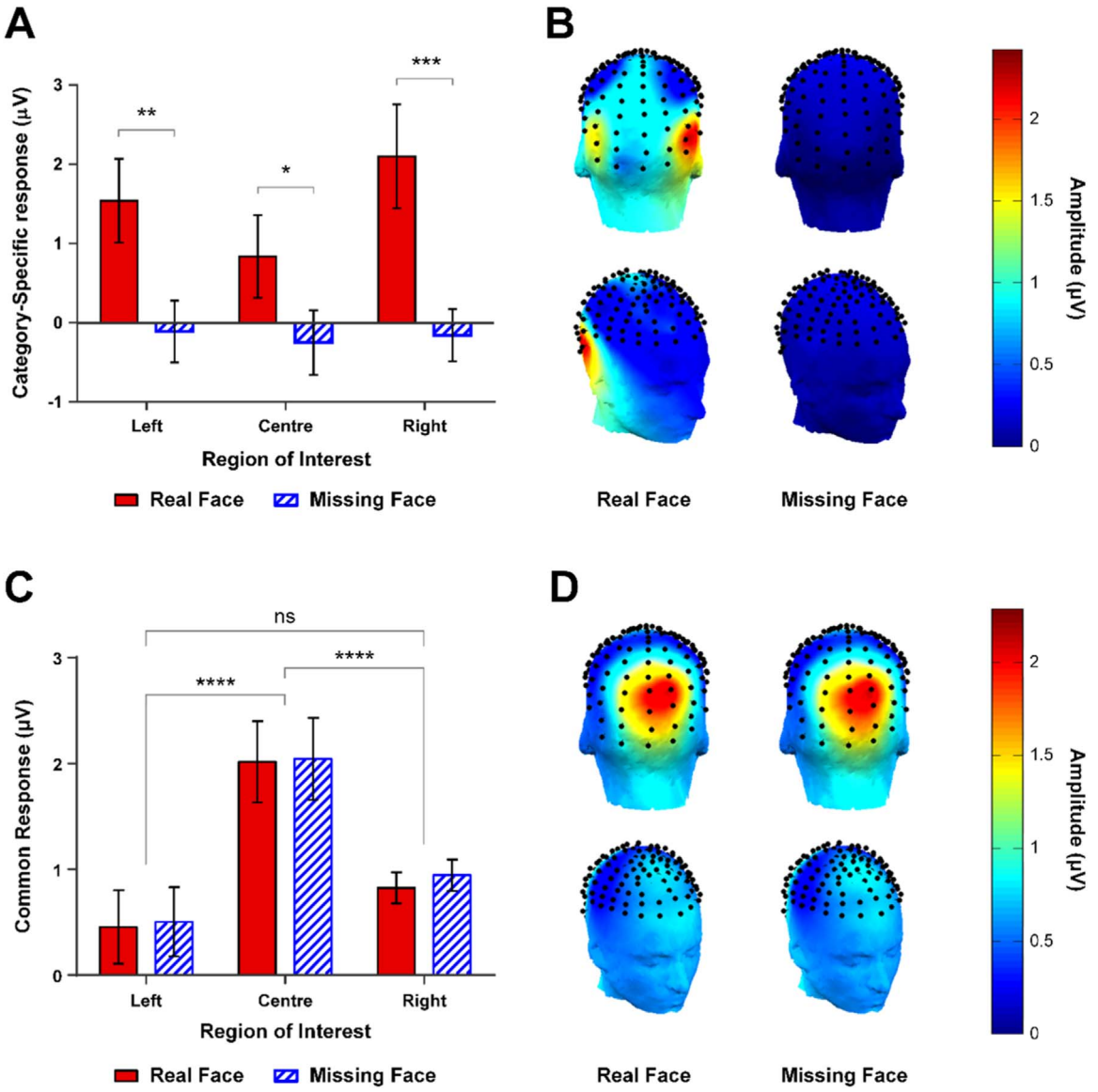


Fig. 11. Sum of baseline-corrected amplitudes for the false-sequence data in Expt. 2. **(A)** The category-selective response was strongly modulated by both ROI and false-sequence type. In all ROIs, real faces generated a clear differential response, where "missing faces" did not. **(C)** In contrast, the common response did not differ with false-sequence type, but was strongly modulated by ROI. Significance codes: **** $p < .0001$; *** $p < .001$; ** $p < .01$; * $p < .05$. Error bars are within-subjects 95% CIs, overlap should not be interpreted by eye (Cumming and Finch, 2005). **(B)** & **(D)** scalp topographies corresponding to the category-selective and common responses, shown for real and missing faces.

As a secondary point, it is worth noting that even though the overall periodicity of the faces in these sequences was degraded (i.e., 10% of periodic faces replaced by objects), our initial frequency-domain analysis of the intact sequences still revealed a clear category-selective response. This is an encouraging finding, since it suggests the category-selective response measured with rapid continuous visual stimulation designs can tolerate imperfection in the periodicity of critical category exemplars (the exact amount of degradation remains an empirical question). This advantage, along with the high signal-to-noise ratio offered by the technique, makes FPVS an ideal method for testing participants who may have particularly noisy EEG signals or difficulty following instructions about tasks or blinks, such as young infants (see de Heering and Rossion, 2015) or clinical populations.

10. General discussion

The goal of our study was to establish whether the category-selective signal elicited by the periodic appearance of critical category exemplars in a dynamic visual stimulation stream is generated or modulated by participants' temporal expectations about those exemplars. To this end, we examined *i)* whether temporally predictable and unpredictable faces embedded in a continuous stream of objects elicit different neural responses (Expt. 1), and *ii)* whether there is an identifiable neural response to a violation of the faces' temporal predictability (Expt. 2).

The research here establishes several important results. First, the category-selective response for faces embedded in a dynamic stream of objects does not vary as a function of the faces' temporal predictability when participants are engaged in an orthogonal task. In Expt. 1, frequency- and time-domain analyses indicated that both the magnitude of the category-selective response and its spatiotemporal unfolding are entirely comparable for periodic and nonperiodic faces. That the response to both face types was indistinguishable at the scalp level speaks directly to the concern that the category-selective response in FPVS designs may be driven in part by the temporal predictability of the critical stimuli. Our results emphatically suggest that this is not the case. Importantly, since even the relatively late category-selective components were insensitive to temporal predictability (e.g., the P3-

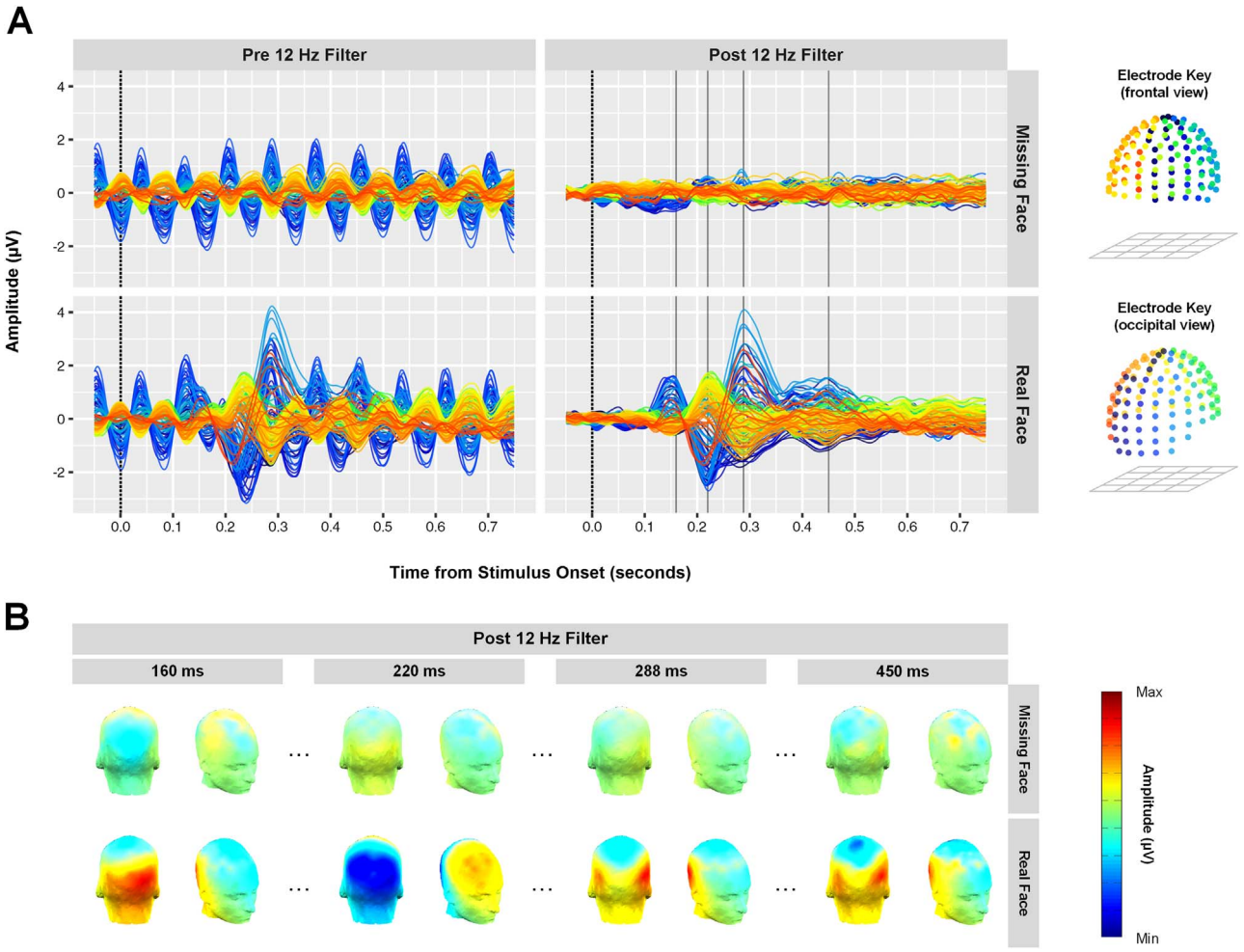


Fig. 12. Time-domain responses in Expt. 2. **(A)** Conditional mean waveforms time-locked to missing face (top row) and real face (bottom row) onset. Left column: the evoked response before selectively removing the 12 Hz response common to both faces and objects. Right column: The differential response to faces after removing the 12 Hz response. Only real faces gave rise to a clear differential (i.e., face-selective) response, comprised of the same four spatiotemporal components seen in Expt. 1. **(B)** Group-level scalp topographies for missing and real faces corresponding to the four components of the face response (160 ms, 220 ms, 288 ms, & 450 ms). Amplitude scales are fixed for condition pairs, but vary across time points.

face, peaking around 300 ms), our data imply that the entire differential response to faces in FPVS designs reflects a categorization response, rather than a response due to temporal expectancy. A second major finding is that rare omissions of highly temporally predictable faces do not evoke a differential response. In Expt. 2, we showed that the response evoked by object images that unexpectedly replace temporally predictable faces is no different to that evoked by other object images about which participants have no temporal expectations. In other words, any temporal expectation induced by the periodicity of faces was not sufficient to generate a category-selective response in the absence of a true face stimulus.

That temporal predictability did not appear to modulate or drive the category-selective signal in our experiments is somewhat surprising given the wealth of temporal expectation effects documented in the literature (Ariga and Yokosawa, 2008; Mathewson et al., 2010; Kok et al., 2012a, 2014; Rohenkohl et al., 2012; Cravo et al., 2013; Breska and Deouell, 2014; Morillon et al., 2016). Before considering possible explanations for these divergent findings, however, we must first rule out the possibility that the lack of temporal periodicity effects here is merely Type II error. We do not believe this to be the case, as in Expt. 1, we know from the initial frequency analysis that an *N* of 18 yields sufficient power to detect differences between conditions (see Fig. 3). Moreover, Bayes factor analysis of these data indicated the null model accounted substantially better for the data than a model including face periodicity. Note that a floor effect does not account for the lack of difference between conditions either – there was a highly significant face categorization response in both conditions in every individual participant's data. In Expt. 2, the minimal response to missing faces was present not only at the group level, but also for each individual participant (Fig. S2 in Supplemental materials). As such, it is not the case that we did not have enough power to detect a significant response signal for missing faces at the group level. Rather, this signal was simply not present for any participant. Overall, given the convergent results of our several analysis methods, as well as the additional precautions we have taken regarding drawing inferences from *p*-values in support of the null hypothesis (Dienes, 2014), we believe the null results we report here can be relied upon.

As suggested above, one possibility is that the periodicity of faces in our two experiments *did* elicit temporal expectations in participants, but that the high saliency of face stimuli for the visual system (Hershler and Hochstein, 2005; Crouzet et al., 2010) results in an evoked response that is simply too robust to be modulated by temporal expectation. On this possibility, if we had used a less potent critical stimulus category (e.g., pictures of houses or body parts, Jacques et al., 2016), or degraded/ambiguous face stimuli (Zhang et al., 2008), we might well have observed an effect of temporal expectation. While we cannot rule out this 'robust-to-expectation-effects' explanation on the basis of the current data (as the critical category was faces in both experiments), it is somewhat undermined by the fact that other forms of top-down expectation (e.g., category level expectation, such as

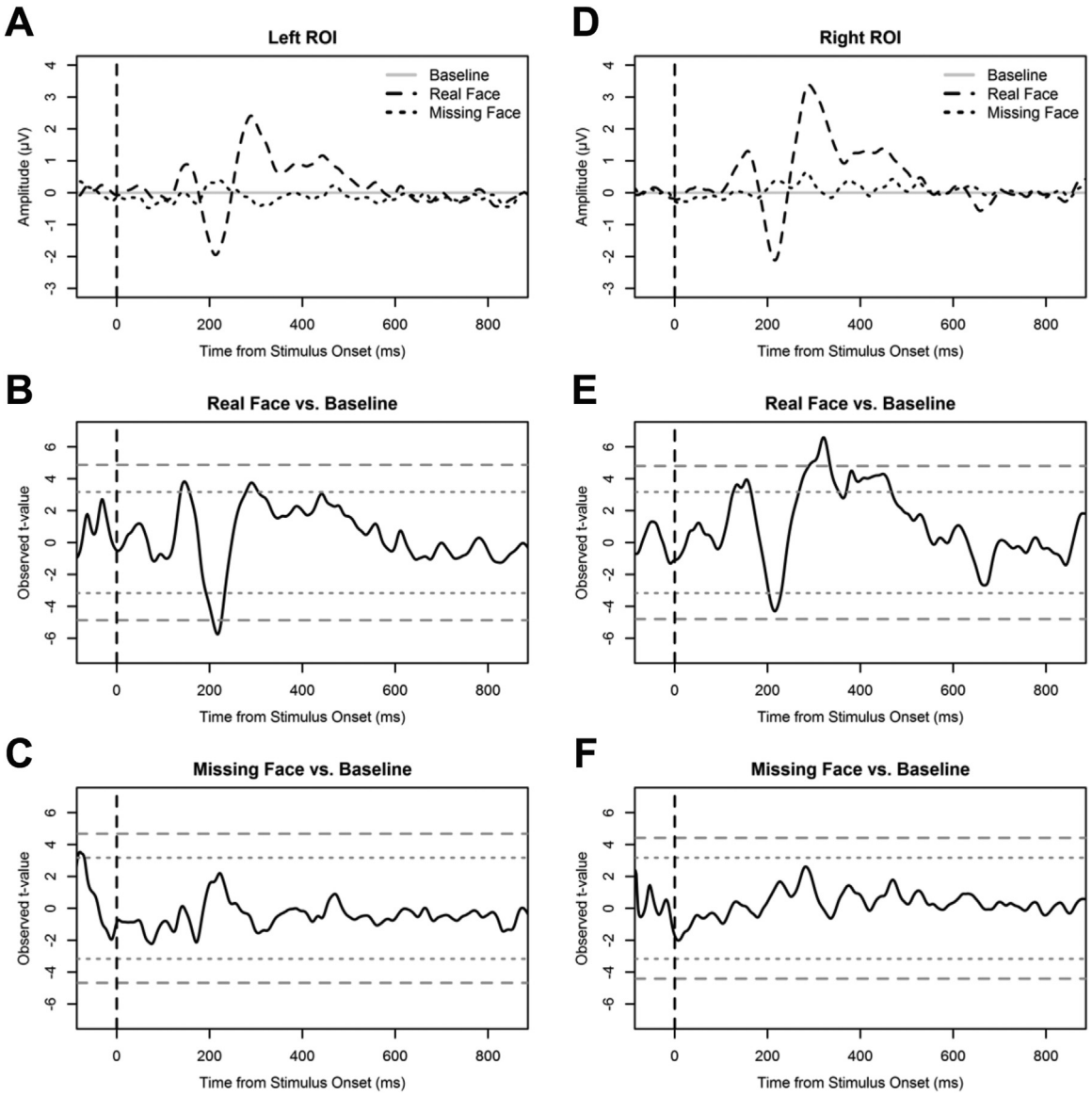


Fig. 13. Time-domain analysis for the left and right ROIs in Expt. 2. Left Column: **(A)** Conditional mean amplitude in the left ROI after selectively removing the 12 Hz response, shown as a function of time from stimulus onset (vertical dashed line). **(B)** Left ROI analysis of the 'real face' response (real face vs. baseline), showing observed *t*-values for each time point. **(C)** Left ROI analysis of the missing face response (missing face vs. baseline), showing observed *t*-values for each time point. **(D)**, **(E)**, & **(F)** are the same panels for the right ROI. Horizontal dashed lines are significance criteria (long dash = $p < .05$, permutation corrected; short dash = $p < .01$, uncorrected). In both ROIs, only observed *t*-values in the real face condition met either criterion for significance, suggesting the response in the missing face condition did not differ from baseline at any point following stimulus onset.

expecting to see a face instead of a house), have been shown to be perfectly capable of modulating the neural response to face stimuli (Gregory, 1970; Esterman and Yantis, 2009; Puri et al., 2009; Egner et al., 2010; Jiang et al., 2015).

An alternative possibility is that face periodicity in our sequences simply did not engender temporal expectations about faces in our participants. Although a number of studies have shown that rhythmic visual stimulation can induce temporal expectations in observers (Ariga and Yokosawa, 2008; Mathewson et al., 2010; Rohenkohl et al., 2012; Cravo et al., 2013; Breska and Deouell, 2014), in our case the perceptual load imposed by the dynamic visual stimulation may have undermined participants' ability to form reliable and accurate temporal expectations about the faces in the sequences. The FPVS approach places the visual system under significant strain – observers see a very large number of images (e.g., 720 images in a single minute) at an extremely high frequency of presentation (e.g., 12 images per second), with each image viewed for just 83.33 ms before being replaced. Within this dynamic and rapid stimulation, the critical periodicity participants must detect is in fact an *embedded* periodicity (i.e., the face presentation frequency at 0.67 Hz, not the common frequency at 12 Hz). Where it is perhaps relatively easy for participants to perceive the rhythmic 'beat' (Rohenkohl et al., 2012) of the common frequency as each new image appears, it may be comparatively harder to perceive the beat of the embedded face instances. This remains an open question at this stage, as we are not aware of any study thus far that has examined temporal expectation effects for embedded visual periodicities. Further to this point, we note that the face stimuli used here were not typical, full frontal exemplars of uniform size, but by design varied widely in terms of viewing angle, face size, lighting, background, etc. As such, any temporal expectations about these faces must necessarily rely on a high-level, category-based stimulus template, rather than one based on low-level commonalities between face exemplars. These temporal and image-based complexities distinguish our study from existing rhythmic visual stimulation studies, which have typically used slower presentation rates and/or shorter, less complex visual stimulations (Mathewson et al., 2010; Rohenkohl et al., 2012; Cravo et al., 2013; Breska and

Deouell, 2014). Where periodicity appears to engender temporal expectations under these simplified conditions, we have shown here that it does not do so when visual stimulation is complex (i.e., highly variable natural images), dynamic, and continuous – conditions which emulate the processing strain imposed on the visual system by real world vision (e.g., when a moving observer makes eye-movements in the context of a dynamic visual scene).

Still another consideration is whether the impact of temporal predictability on stimulus processing might vary with the task-relevance of that stimulus. To the best of our knowledge, previous studies that have demonstrated modulation of perceptual processing by rhythmic temporal expectations have used behaviorally relevant target items (Ariga and Yokosawa, 2008; Rohenkohl et al., 2012; Cravo et al., 2013; Breska and Deouell, 2014; Morillon et al., 2016). That is, the stimulus for which an effect of temporal expectation was observed was one that participants actually responded to in some way. In contrast, the faces in our experiments were never task-relevant for participants, who we instructed to monitor the central fixation cross for color changes. Note that we are not suggesting that task-irrelevant stimuli cannot engender temporal expectations in participants – indeed, several studies have already documented this (e.g., den Ouden et al., 2009; Alink et al., 2010) – but that perhaps in order for temporal expectations to facilitate perceptual processing, the critical stimuli must be attended to in some way. In the present experiments, we deliberately imposed an orthogonal task on the stimulation sequence, so as not to confound expectation and attention related effects (cf. Summerfield and Egner, 2009). Indeed, the FPVS approach has most often been employed this way to yield an implicit index of perceptual categorization (e.g., Rossion et al., 2015; Retter and Rossion, 2016). However, if participants were required to judge some aspect of periodic faces explicitly (e.g., "respond when you see a female face"), we might well have observed an effect of temporal predictability on the category-selective response (Kok et al., 2012b).

Taken together, the findings here support the claim that the index of perceptual categorization yielded by dynamic visual stimulation approaches (e.g., Rossion et al., 2015; Jonas and Rossion, 2016; Retter and Rossion, 2016) is immune to the temporal predictability. Importantly, this finding not only validates the use of FPVS as an objective tool with which to operationalize perceptual categorization processes, but also has important implications for understanding human perceptual categorization in a rapidly changing (i.e., dynamic) visual scene. Furthermore, these data undermine a predictive coding interpretation of category-change detection in the human brain (Rao and Ballard, 1999; Friston, 2005; Alink et al., 2010; Kok et al., 2012a), in that temporally predictable faces in our study were not associated with reduced neural activity (due to their 'redundancy'), or a sharpened sensory representation (due to noise suppression) (Kok et al., 2012a). That top-down temporal expectations do not facilitate sensory processing in the context of a dynamic continuous scene points to the interesting possibility that predictive mechanisms are not automatic, but rather subject to the rate at which visual stimulation changes. In this way, this work highlights a broader need to test theoretical predictions under more ecologically relevant conditions, which place the visual system under a similar strain to that imposed by natural vision.

11. Conclusion

The research reported here establishes that category-selective neural responses elicited in a dynamic visual stream are immune to the temporal predictability of the critical category exemplars, at least when *i)* visual stimuli are highly variable in viewing conditions, and *ii)* participants are engaged in an orthogonal task that does not explicitly highlight the temporal predictability of the critical stimulus. Under such conditions, the category-selective response yielded by these designs can be taken as a relatively pure index of perceptual categorization.

Acknowledgements

The authors are grateful to T. Retter for her helpful insights and expertise in developing the false-sequencing data analysis. This work was supported in part by the European Research Council to BR [Grant no: facessvep 284025] and cofunded by the University of Louvain and the Marie Curie Actions of the European Commission [Grant no: F211800012] to GLQ.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at http://dx.doi.org/10.1016/j.neuropsychologia.2017.08.010.

References

Alink, A., Schwiedrzik, C.M., Kohler, A., Singer, W., Muckli, L., 2010. Stimulus predictability reduces responses in primary visual cortex. J. Neurosci. 30, 2960–2966.
Alonso-Prieto, E., Belle, G.V., Liu-Shuang, J., Norcia, A.M., Rossion, B., 2013. The 6 Hz fundamental stimulation frequency rate for individual face discrimination in the right occipito-temporal cortex. Neuropsychologia 51, 2863–2875.
Ariga, A., Yokosawa, K., 2008. Attentional awakening: gradual modulation of temporal attention in rapid serial visual presentation. Psychol. Res. 72, 192–202.
Bai, O., Lin, P., Vorbach, S., Li, J., Furlani, S., Hallett, M., 2007. Exploration of computational methods for classification of movement intention during human voluntary movement from single trial EEG. Clin. Neurophysiol. 118, 2637–2655.
Blair, R.C., Karniski, W., 1993. An alternative method for significance testing of waveform difference potentials. Psychophysiology 30, 518–524.
Blankertz, B., Lemm, S., Treder, M., Haufe, S., Müller, K.-R., 2011. Single-trial analysis and classification of ERP components—a tutorial. NeuroImage 56, 814–825.
Bode, S., Stahl, J., 2014. Predicting errors from patterns of event-related potentials preceding an overt response. Biol. Psychol. 103, 357–369.
Bode, S., Sewell, D.K., Lilburn, S., Forte, J.D., Smith, P.L., Stahl, J., 2012. Predicting perceptual decision biases from early brain activity. J. Neurosci. 32, 12488–12498.
Bode, S., Bennett, D., Feuerriegel, D., Alday, P., 2016. The Decision Decoding Toolbox – A Multivariate Pattern Analysis Toolbox for Event-Related Potentials. http://dx.doi.org/10.5281/zenodo.48143.
Breska, A., Deouell, L.Y., 2014. Automatic bias of temporal expectations following temporally regular input independently of high-level temporal expectation. J. Cogn. Neurosci. 26, 1555–1571.
Cavanagh, P., Leclerc, Y.G., 1989. Shape from shadows. J. Exp. Psychol. Hum. 15, 3.
Chang, C.-C., Lin, C.-J., 2011. LIBSVM: a library for support vector machines. ACM Trans. Intell. Syst. Technol. 2, 27.
Cravo, A.M., Rohenkohl, G., Wyart, V., Nobre, A.C., 2013. Temporal expectation enhances contrast sensitivity by phase entrainment of low-frequency oscillations in visual cortex. J. Neurosci. 33, 4002–4010.
Crouzet, S.M., Thorpe, S.J., 2011. Low-level cues and ultra-fast face detection. Front. Psychol. 2, 342.
Crouzet, S.M., Kirchner, H., Thorpe, S.J., 2010. Fast saccades toward faces: face detection in just 100 ms. J. Vis. 10 (16-16).
Cumming, G., Finch, S., 2005. Inference by eye: confidence intervals and how to read pictures of data. Am. Psychol. 60, 170–180.
Das, K., Giesbrecht, B., Eckstein, M.P., 2010. Predicting variations of perceptual performance across individuals from neural activity using pattern classifiers. Neuroimage 51, 1425–1437.
Dienes, Z., 2014. Using Bayes to get the most out of non-significant results. Front. Psychol. 5, 781.
Egner, T., Monti, J.M., Summerfield, C., 2010. Expectation and surprise determine neural population responses in the ventral visual stream. J. Neurosci. 30, 16601–16608.
Esterman, M., Yantis, S., 2009. Perceptual expectation evokes category-selective cortical activity. Cereb. Cortex.
Finkbeiner, M., Friedman, J., 2011. The flexibility of nonconsciously deployed cognitive processes: evidence from masked congruence priming. PLoS One 6, e17095.
Finkbeiner, M., Palermo, R., 2009. The role of spatial attention in nonconscious processing: a comparison of face and nonface stimuli. Psychol. Sci. 20, 42–51.
Friston, K., 2005. A theory of cortical responses. Philos. Trans. R. Soc. Lond. B Biol. Sci. 360, 815–836.
Garrido, M.I., Kilner, J.M., Stephan, K.E., Friston, K.J., 2009. The mismatch negativity: a review of underlying mechanisms. Clin. Neurophysiol. 120, 453–463.
Gregory, R.L., 1970. The Intelligent Eye. McGraw-Hill, New York.
de Heering, A., Rossion, B., 2015. Rapid categorization of natural face images in the infant right hemisphere. eLife 4, e06564.
Hershler, O., Hochstein, S., 2005. At first sight: a high-level pop out effect for faces. Vis. Res. 45, 1707–1724.
Hyvarinen, A., Oja, E., 2000. Independent component analysis: algorithms and applications. Neural Netw. 13, 411–430.
Jacques, C., Retter, T.L., Rossion, B., 2016. A single glance at a face generates larger and qualitatively different category-selective spatiotemporal signatures than other ecologically-relevant categories in the human brain. NeuroImage 137, 21–33.

Jeffreys, H., 1939/1961. The Theory of Probability. Oxford University Press, Oxford, England.
Jiang, F., Badler, J.B., Righi, G., Rossion, B., 2015. Category search speeds up face-selective fMRI responses in a non-hierarchical cortical face network. Cortex 66, 69–80.
Jonas, J., Rossion, B., 2016. Beyond the core face-processing network: intracerebral stimulation of a face-selective area in the right anterior fusiform gyrus elicits transient prosopagnosia. J. Vis. 16 (385-385).
Jones, M.R., 1976. Time, our lost dimension: toward a new theory of perception, attention, and memory. Psychol. Rev. 83, 323–355.
Kimura, M., Schröger, E., Czigler, I., 2011. Visual mismatch negativity and its importance in visual cognitive sciences. NeuroReport 22, 669–673.
Kok, P., Jehee, J.F., de Lange, F.P., 2012a. Less is more: expectation sharpens representations in the primary visual cortex. Neuron 75, 265–270.
Kok, P., Rahnev, D.A., Jehee, J.F.M., Lau, H.C., de Lange, F.P., 2012b. Attention reverses the effect of prediction in silencing sensory signals. Cereb. Cortex 22, 2197–2206.
Kok, P., Failing, M.F., de Lange, F.P., 2014. Prior expectations evoke stimulus templates in the primary visual cortex. J. Cogn. Neurosci. 26, 1546–1554.
Lee, T.S., Mumford, D., 2003. Hierarchical Bayesian inference in the visual cortex. J. Opt. Soc. Am. A Opt. Image Sci. Vis. 20, 1434–1448.
Lieder, F., Stephan, K.E., Daunizeau, J., Garrido, M.I., Friston, K.J., 2013. A neurocomputational model of the mismatch negativity. PLoS Comput. Biol. 9, e1003288.
Liu-Shuang, J., Norcia, A.M., Rossion, B., 2014. An objective index of individual face discrimination in the right occipito-temporal cortex by means of fast periodic oddball stimulation. Neuropsychologia 52, 57–72.
Lochy, A., Van Belle, G., Rossion, B., 2015. A robust index of lexical representation in the left occipito-temporal cortex as evidenced by EEG responses to fast periodic visual stimulation. Neuropsychologia 66, 18–31.
Lochy, A., Van Reybroeck, M., Rossion, B., 2016. Left cortical specialization for visual letter strings predicts rudimentary knowledge of letter-sound association in preschoolers. Proc. Natl. Acad. Sci. USA 113, 8544–8549.
Luck, S.J., 2005. An Introduction to the Event-Related Potential Technique. MIT, Cambridge, Mass London.
Mathewson, K.E., Fabiani, M., Gratton, G., Beck, D.M., Lleras, A., 2010. Rescuing stimuli from invisibility: Inducing a momentary release from visual masking with pre-target entrainment. Cognition 115, 186–191.
McAuley, J.D., Jones, M.R., 2003. Modeling effects of rhythmic context on perceived duration: a comparison of interval and entrainment approaches to short-interval timing. J. Exp. Psychol. Hum. 29, 1102–1125.
Mooney, C.M., 1957. Age in the development of closure ability in children. Can. J. Psychol. 11, 219.
Moore, C., Cavanagh, P., 1998. Recovery of 3D volume from 2-tone images of novel objects. Cognition 67, 45–71.
Morillon, B., Schroeder, C.E., Wyart, V., Arnal, L.H., 2016. Temporal prediction in lieu of periodic stimulation. J. Neurosci. 36, 2342–2347.
Müller, M.M., Andersen, S., Trujillo, N.J., Valdés-Sosa, P., Malinowski, P., Hillyard, S.A., 2006. Feature-selective attention enhances color signals in early visual areas of the human brain. Proc. Natl. Acad. Sci. USA 103, 14250–14254.
Naatanen, R., Astikainen, P., Ruusuvirta, T., Huotilainen, M., 2010. Automatic auditory intelligence: an expression of the sensory-cognitive core of cognitive processes. Brain Res. Rev. 64, 123–136.
Norcia, A.M., Appelbaum, L.G., Ales, J.M., Cottereau, B.R., Rossion, B., 2015. The steady-state visual evoked potential in vision research: a review. J. Vis. 15 (4-4).
den Ouden, H.E., Friston, K.J., Daw, N.D., McIntosh, A.R., Stephan, K.E., 2009. A dual role for prediction error in associative learning. Cereb. Cortex 19, 1175–1185.
Pieszek, M., Widmann, A., Gruber, T., Schröger, E., 2013. The human brain maintains contradictory and redundant auditory sensory predictions. PLoS One 8, e53634.
Potter, M.C., Levy, E.I., 1969. Recognition memory for a rapid sequence of pictures. J. Exp. Psychol. 81, 10.
Puri, A.M., Wojciulik, E., Ranganath, C., 2009. Category expectation modulates baseline and stimulus-evoked activity in human inferotemporal cortex. Brain Res. 1301, 89–99.
Quek, G.L., Finkbeiner, M., 2013. Spatial and temporal attention modulate the early stages of face processing: behavioural evidence from a reaching paradigm. PLoS One 8, e57365.
Rajendran, V.G., Teki, S., 2016. Periodicity versus prediction in sensory perception. J. Neurosci. 36, 7343–7345.
Rao, R.P., Ballard, D.H., 1999. Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. Nat. Neurosci. 2, 79–87.
Reddy, L., Wilken, P., Koch, C., 2004. Face-gender discrimination is possible in the near-absence of attention. J. Vis. 4, 106–117.
Reddy, L., Reddy, L., Koch, C., 2006. Face identification in the near-absence of focal attention. Vis. Res. 46, 2336–2343.
Retter, T.L., Rossion, B., 2016. Uncovering the neural magnitude and spatiotemporal dynamics of natural image categorization in a fast visual stream. Neuropsychologia 2016 (91), 9–28.
Rohenkohl, G., Cravo, A.M., Wyart, V., Nobre, A.C., 2012. Temporal expectation improves the quality of sensory information. J. Neurosci. 32, 8424–8428.
Rossion, B., 2014. Understanding individual face discrimination by means of fast periodic visual stimulation. Exp. Brain Res. 232, 1599–1621.
Rossion, B., Boremanse, A., 2011. Robust sensitivity to facial identity in the right human occipito-temporal cortex as revealed by steady-state visual-evoked potentials. J. Vis. 11 (2).
Rossion, B., Prieto, E.A., Boremanse, A., Kuefner, D., Van Belle, G., 2012. A steady-state visual evoked potential approach to individual face perception: effect of inversion, contrast-reversal and temporal dynamics. Neuroimage 63, 1585–1600.
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. J. Vis. 15 (15.11.18).
Rousselet, G.A., Mace, M.J., Fabre-Thorpe, M., 2003. Is it an animal? Is it a human face? Fast processing in upright and inverted natural scenes. J. Vis. 3, 440–455.
Stefanics, G., Kimura, M., Czigler, I., 2011. Visual mismatch negativity reveals automatic detection of sequential regularity violation. Front. Hum. Neurosci. 5, 46.
Summerfield, C., Egner, T., 2009. Expectation (and attention) in visual cognition. Trends Cogn. Sci. 13, 403–409.
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, 229–233.