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

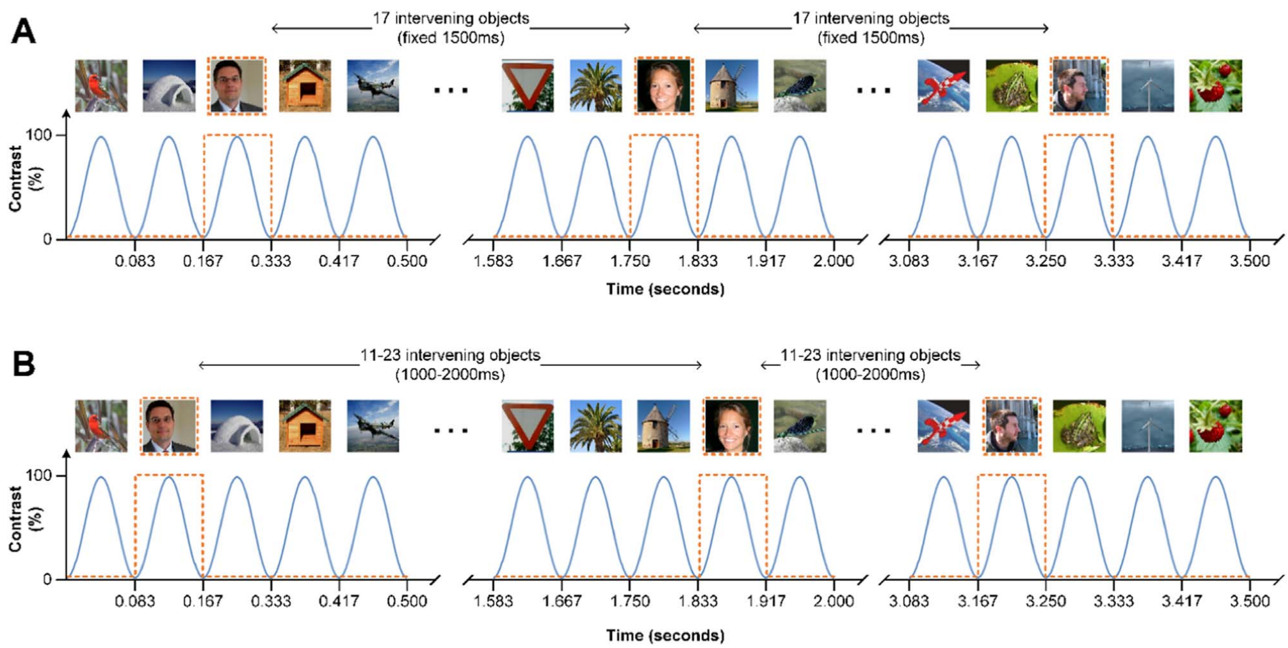

Fig. 1. Expt. 1 sequence design. We presented images at a rate of exactly 12 Hz by sinusoidally modulating the contrast of each from 0–100-0% (blue solid lines). In both conditions, the 90 s sequence contained 60 natural face images and numerous natural object images (e.g., vehicles, animals, buildings, 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;