Vision Research 145 (2018) 11–20

Contents lists available at ScienceDirect

Vision Research

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

Dissociable effects of inter-stimulus interval and presentation duration on rapid face categorization

Talia L. Retter[a,b,*], Fang Jiang[b], Michael A. Webster[b], Bruno Rossion[a,c]

[a] Psychological Sciences Research Institute, Institute of Neuroscience, University of Louvain, Belgium
[b] Department of Psychology and Graduate Program in Integrative Neuroscience, University of Nevada, Reno, USA
[c] Neurology Unit, Centre Hospitalier Regional Universitaire (CHRU) de Nancy, F-54000 Nancy, France

ARTICLE INFO

Keywords:
Vision
EEG
Face perception
Duration
Masking
Frequency tagging

ABSTRACT

Fast periodic visual stimulation combined with electroencephalography (FPVS-EEG) has unique sensitivity and objectivity in measuring rapid visual categorization processes. It constrains image processing time by presenting stimuli rapidly through brief stimulus presentation durations and short inter-stimulus intervals. However, the selective impact of these temporal parameters on visual categorization is largely unknown. Here, we presented natural images of objects at a rate of 10 or 20 per second (10 or 20 Hz), with faces appearing once per second (1 Hz), leading to two distinct frequency-tagged EEG responses. Twelve observers were tested with three squarewave image presentation conditions: 1) with an ISI, a traditional 50% duty cycle at 10 Hz (50-ms stimulus duration separated by a 50-ms ISI); 2) removing the ISI and matching the rate, a 100% duty cycle at 10 Hz (100-ms duration with 0-ms ISI); 3) removing the ISI and matching the stimulus presentation duration, a 100% duty cycle at 20 Hz (50-ms duration with 0-ms ISI). The face categorization response was significantly decreased in the 20 Hz 100% condition. The conditions at 10 Hz showed similar face-categorization responses, peaking maximally over the right occipito-temporal (ROT) cortex. However, the onset of the 10 Hz 100% response was delayed by about 20 ms over the ROT region relative to the 10 Hz 50% condition, likely due to immediate forward-masking by preceding images. Taken together, these results help to interpret how the FPVS-EEG paradigm sets temporal constraints on visual image categorization.

1. Introduction

Fast periodic visual stimulation with electroencephalography (FPVS-EEG) has advantages in sensitivity and objectivity for measuring rapid visual categorization processes through the insertion of within-category (face) images as a proportion of rapidly presented across-category (object) images (e.g., Rossion, Torfs, Jacques, & Liu-Shuang, 2015; Jacques, Retter, & Rossion, 2016; Jonas et al., 2016; Retter & Rossion, 2016). In the example illustrated in Fig. 1A, natural images are presented at a fixed periodic rate, i.e., 10 images per second with faces appearing as every 1 out of 10 images, leading to frequency-tagged EEG responses (also known as "steady-state visual evoked potentials", SSVEPs; for a general review, see Norcia, Appelbaum, Ales, Cottereau, & Rossion, 2015) at 10 Hz for image presentation and 1 Hz for face-selective responses. By using highly variable natural face images, as well as a wide variety of natural object images, the paradigm is able to capture both generalization across within-category face exemplars and discrimination of face-selective vs. generic visual responses (Rossion et al., 2015).

FPVS-EEG is not only defined by the periodic presentation of visual stimuli: it is also defined by its relatively *fast* rate of stimulation (i.e., compared to standard EEG or behavioral studies). A relatively fast rate, e.g., 4–12 Hz, is valuable because it constrains neural responses to limited time windows. This borrows from the logic of rapid serial visual presentation (RSVP), in which stimuli are serially presented, sometimes in the range of eye fixation rates or faster, with participants subsequently providing responses to the images they perceived or remembered (Potter, 2012; Potter & Levy, 1969; Potter, Wyble, Hagmann, & McCourt, 2014). The brief image duration, as well as perceptual backward-masking from sequential stimuli (typically shown without any inter-stimulus interval) used in RSVP temporally limits the availability of visual information about a stimulus category (Potter, 2012). Consequently, RSVP has been used to probe the contributions of attention and memory to visual perception at a behavioral level. RSVP has also been applied to single neuron recordings to investigate the effects of forward and backward masking, as well as the temporal limits of image processing in the brain (Keysers & Perrett, 2002; Keysers, Xiao, Foldiak, & Perrett, 2005; Keysers, Xiao, Földiák, & Perrett, 2001).

* Corresponding author at: 1664 N Virginia St., Mail Stop 296, Reno, Nevada 89557, USA.
E-mail address: tlretter@nevada.unr.edu (T.L. Retter).

https://doi.org/10.1016/j.visres.2018.02.009
Received 20 September 2017; Received in revised form 7 February 2018; Accepted 8 February 2018
Available online 12 April 2018
0042-6989/

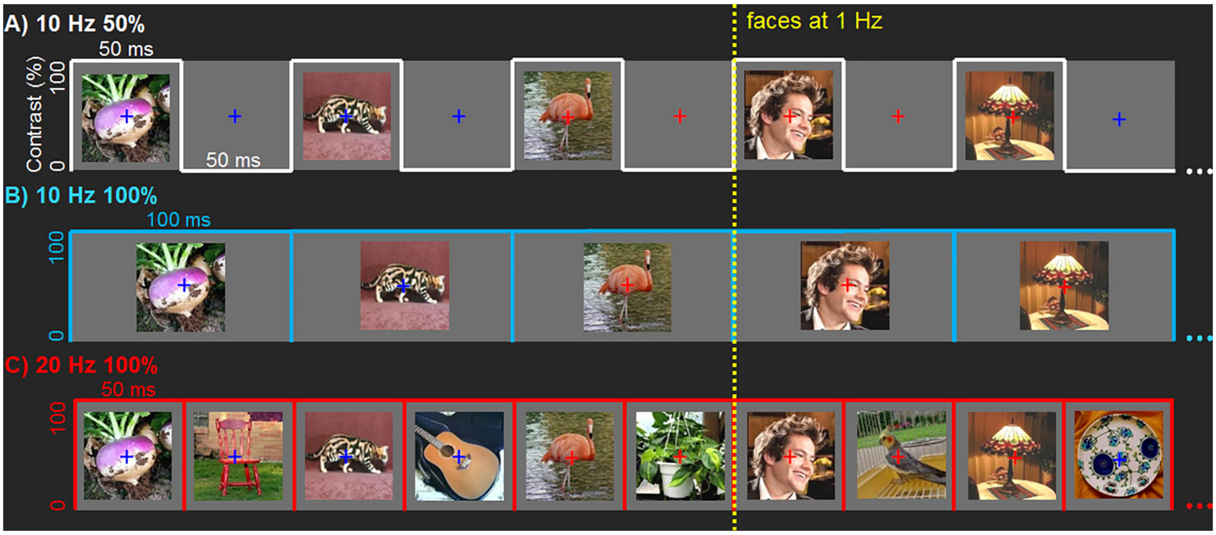

Fig. 1. Experimental paradigm and conditions. In all stimulus presentation modes, images of natural objects are presented in a random order without repetition, and face stimuli appear every 1 s, i.e., at a rate of 1 Hz. An excerpt of only 500 ms is depicted here for each testing sequence, including an example of the brief fixation cross color change, providing cues for an orthogonal behavioral task. A) 10 Hz 50%: Stimuli are presented at 10 Hz with a 50% duty cycle, i.e., "on" for 50 ms at full luminance contrast and "off" for an ISI of 50 ms at 0% contrast. B) 10 Hz 100%: Stimuli are presented at 10 Hz again, but with a 100% duty cycle, staying on for 100 ms with no ISI. C) 20 Hz 100%: Stimuli are presented at 20 Hz with a 100% duty cycle, appearing on for 50 ms with no ISI, matching the stimulus presentation duration of the first condition but the 0-ms ISI of the second condition.

Building on these techniques, FPVS-EEG presents longer streams of visual images with embedded frequency-tagged categories, allowing for a category-specific neural response to be characterized and quantified (Retter & Rossion, 2016).

The rate of image presentation can be controlled with two factors: 1) the duration of stimulus presentation and 2) the inter-stimulus interval (ISI). For different experimental purposes, stimuli may be presented with or without an ISI, e.g., with an ISI to prevent image after-effects or without an ISI to maximize forward and backward masking, so as to limit visual information. Indeed, removing the ISI may have implications on the degree of backward and forward masking of each stimulus (Crawford, 1947; Macknik & Livingstone, 1998). With relatively short ISIs (below about 90 or 100 ms), visual short term memory may serve to "fill-in" the gaps between stimuli (Potter, 2012), leading to no differences reported in human image detection performance or in monkey superior-temporal sulcus neuron firing rates (Keysers & Perrett, 2002; see also Rolls & Tovée, 1994). Subtle differences may be observed, however, in the onset of neural responses to stimuli presented with or without an ISI. Specifically, there was a small, uninvestigated delay of responses without an ISI in a previous study (see Fig. 3A of Keysers & Perrett, 2002; also in Fig. 3 of Keysers et al., 2005). This might be due to transient off and/or sustained responses produced from forward masking (i.e., from previously presented images) overlapping with and suppressing subsequent image responses (e.g., Breitmeyer, Levi, & Harwerth, 1981; Duysens, Orban, Cremieux, & Maes, 1985; Keysers et al., 2005; Macknik & Livingstone, 1998; Ogmen, Breitmeyer, & Melvin, 2003).

In this study we asked how the neural responses captured by the FPVS face categorization paradigm depend on the temporal properties of the stimulus sequence. To separate the effect of ISI from stimulus duration on rapid face categorization, we compared three squarewave[1] stimulation modes. The first of these was chosen to closely replicate parameters used in previous studies (e.g., Dzhelyova et al., 2016; Retter & Rossion, 2016), in which stimuli were presented at 10 times a second (10 Hz) with a 50% duty cycle (each image shown for 50 ms and followed by a 50 ms ISI) (Fig. 1A). In the second condition, we removed the ISI but used the same presentation rate. That is, we presented stimuli at 10 Hz but with a 100% duty cycle, presenting images for 100 ms with no ISI (Fig. 1B). In the third condition, we also removed the ISI but used the original presentation duration, so that the images were shown at 20 Hz with a 100% duty cycle, displaying images for 50 ms with no ISI (Fig. 1C). In all conditions, faces appeared at 1 Hz, i.e., with a face stimulus shown every 1 s, and we asked how the 1 Hz response to faces (reflecting face categorization) depended on the temporal properties of the image stream.

Specifically, by comparing the first two conditions, the effect of the ISI (50 vs. 0 ms) may be investigated while controlling the presentation rate at 10 Hz. By comparing the first and third conditions, the effect of ISI (again 50 vs. 0 ms) may be investigated while controlling the image presentation duration at 50 ms. Finally, by comparing the second and third conditions, the effect of stimulus presentation duration (100 vs. 50 ms) may be investigated when the ISI is not present. Practically, it is important to understand the effects of different presentation modes when designing experiments. For example, an ISI may be desired for limiting after-effects (e.g., color or face-related) or apparent motion across sequential stimulus presentations; however, an ISI may be undesired for testing the temporal constraints that forward and backward masking impose on (e.g., the speed of) perception. Thus, one goal was to provide information about the appropriateness of different presentation modes for future studies, as well as to provide baseline data for comparing across past and future studies using different image presentation modes. Theoretically, we also sought to compare the impacts of an ISI on human population-level neural responses to the effects of ISI previously observed in single unit responses.

2. Materials and Methods

2.1. Participants

Participants included 12 students or employees at the University of

[1] Note that a squarewave presentation was used here in order to identify exactly the presentation duration, whereas the majority of previous FPVS studies applied a sinusoidal modulation of image contrast (e.g., as in Dzhelyova, Jacques, & Rossion, 2016; Jacques, Retter, & Rossion, 2016; Regan, 1966; Retter & Rossion, 2016; Rossion et al., 2015; Silberstein, Schier, Pipingas, et al., 1990; van der Tweel & Verduyn Lunel, 1965; Victor & Zemon, 1985). In sinusoidal presentation, the image display appears smooth and relatively continuous but the contrast level at which each image is identifiable, defining the effective presentation duration, is not always known.