International Journal of Psychophysiology 74 (2009) 69–73

Contents lists available at ScienceDirect

International Journal of Psychophysiology

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

The vase–face illusion seen by the brain: An event-related brain potentials study

Jiang Qiu *, Dongtao Wei, Hong Li, Caiyun Yu, Ting Wang, Qinglin Zhang *

Key Laboratory of Cognition and Personality (SWU), Ministry of Education, Chongqing, China
School of Psychology, Southwest University, Beibei, Chongqing 400715, China

ARTICLE INFO

Article history:
Received 19 February 2009
Received in revised form 23 June 2009
Accepted 20 July 2009
Available online 29 July 2009

Keywords:
The vase–face ambiguous figure
Perceptual reversals
Reversal negativity
Event-related potential (ERP)

ABSTRACT

In the present study, we provided a cue (face or vase) beforehand to make subjects attend the target stimuli in the vase–face ambiguous figure, and compared the spatiotemporal cortical activation patterns underlying face (face–face) or vase (vase–vase) processing to the ambiguous figure using high-density (64-channel) event-related brain potential (ERP) recordings. Scalp ERP analysis found that the anterior N100, P160, N320 and posterior P100 and N160 were elicited by the face–face and vase–vase responses. The results of the ANOVAs showed that the anterior N100 and N320 elicited by the vase–vase response were more negative than the face–face response. The anterior N100 might reflect deployment of attention (conscious effort) to identify the target stimuli (face or vase) in early processing of the ambiguous figure, and the N320 might be the reversal negativity (RN) and was involved in involuntary perceptual reversals (from face to vase or reverse). Moreover, the mean amplitude (a late positive component: LPC) between 350 and 450 ms of the face–face response was larger than the vase–vase response on positive orientation over the front-central scalp regions. This result might support the view that the LPC reflect post-perceptual processing and indicated that the perceptual reversal of the vase–face illusion is influenced by top-down control.

1. Introduction

Bistable stimuli are invaluable tools for the study of the neural basis of perceptual reversals, for example, the famous Rubin vase–face ambiguous figure (see Palmer, 1999; Hulleman and Humphreys, 2004; Pitts et al., 2008). As Eagleman (2001) said, "the interesting property of bistable stimuli is that they can flip back and forth between different perceptions. Although nothing changes on the page, there is more than one way for the visual system to interpret the stimulus, and perceptual reversals occur. The perceptual reversal indicates that cortical processing is an active process that tries to make sense of incoming information". It is well known that the face–vase figure can be interpreted as two faces looking at each other in front of a rectangle, and can also be seen as a vase in front of the same rectangle (Palmer, 1999).

As for now, a longstanding debate exists in the literature concerning the cognitive and neural mechanisms of bistable perception (perceptual reversal). Kornmeier and Bach (2005) had indicated that most hypothetical explanations about the neural processes underlying spontaneous perceptual reversals of ambiguous figures fall into two classes, emphasizing either bottom-up, or top-down factors. Specifically, they summarized, "the bottom-up approach assumes that perceptual reversals are caused by passive adaptation of local neural units during early visual processing (e.g., Toppino and Long, 1987; Brigner and Deni, 1990), while the top-down approach assumes that reversals take place during perceptual interpretation and thus later in the processing hierarchy, near awareness (e.g. Rock et al., 1994; Struber and Stadler, 1999; Long and Toppino, 2004)."

Previous studies using functional magnetic resonance imaging (fMRI) had obtained some important findings about the neural mechanism of the vase–face ambiguous figure perception, for example, some studies have shown that extrastriate regions such as the ventral visual pathway, parietal and frontal regions (Lumer et al., 1998), the fusiform face area and the parahippocampal place area participate in perceptual changes (Tong et al., 1998); Hasson et al. (2001) found that the fusiform gyrus was more active when the vase–face stimulus was biased toward the face by the use of color or texture; in Andrews' et al. (2002) study, their results showed that the responses of regions within the temporal lobe are modulated by selective attention to faces (Wojciulik et al., 1998; O'Craven et al., 1999). Moreover, it is known that ERPs may provide a means to evaluate timing of cognitive processes prior to a response. In the ERP technique, recordings are made of the electrical activity of the brain that is time locked to the presentation of an external stimulus. Thus, ERP data allow for precise statements about the time course of activation during different stages of processing (e.g., low-level visual perception or high-level cognitive control) of the Rubin vase–face ambiguous figure.

Previous ERP studies had found a P300 component or a late positive component (LPC) that is correlated with perceptual reversals, which

* Corresponding authors. School of Psychology, Southwest University, Beibei, Chongqing 400715 China. Tel.: +86 23 6836 7942.
E-mail addresses: qiuj318@swu.edu.cn (J. Qiu), qiuj318@yahoo.com (Q. Zhang).

0167-8760/$ – see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.ijpsycho.2009.07.006

might be involved in top-down voluntary control (e.g., Basar-Eroglu et al., 1993; Isoglu-Alkac et al., 1998; Strüber et al., 2001), while other studies found early ERP components, for example, the face-related P170, reversal positivity as early as 130 ms and reversal negativity at 250 ms, supporting a bottom-up explanation (Kornmeier and Bach, 2004, 2005). Obviously, the neural basis of perceptual reversals remains largely unknown, especially, the issues related to the bottom-up vs. top-down debate still is not resolved. Recently, Pitts et al. (2007) found P1 (120 ms), N1 (175 ms), and the reversal negativity (RN; 170–370 ms) effects related to perceptual reversals, and suggested that these ERP effects as evidence for a role of selective attention in bistable perception. Subsequently, Pitts et al. (2008) required observers to maintain one of three "intentional approaches" to perceive the Necker lattice, (1) try to reverse perception as often as possible, (2) try to stabilize perception for as long as possible, and (3) maintain a passive approach. Their results also indicated that top-down mechanisms can influence perception of bistable figures at least as early as 150 ms post-stimulus onset (indicated by amplitude enhancements of the RN component) and that post-perceptual processing is effected by top-down control (indicated by LPC amplitude differences; Pitts et al., 2008).

In a word, previous studies had found some ERP components related to perception of ambiguous figures, e.g., P1, N1, RN and P300 (LPC). However, as Pitts et al. (2007) indicated, "these findings do not support a strict low-level or high-level theory of multistable perception but, instead, suggest a critical role for perceptual exploration mediated by visual attention mechanisms". That is to say, the specific neural mechanisms that mediate this top-down control remain a topic of debate. In Pitts et al.'s (2008) study, they investigated the attention hypothesis of bistable perception by introducing conditions involving voluntary control (e.g., try to stabilize perception for as long as possible). We thought that this is an effective method to explore how the top-down control (selective attention) influenced the spatiotemporal cortical activation patterns underlying bistable perception. Therefore, in the present study, we tried to present a cue (face or vase) beforehand to make subjects attend the target stimuli in the vase–face ambiguous figure, and compared the spatiotemporal cortical activation patterns underlying face or vase processing to the ambiguous figure using high-density (64-channel) ERP recordings. Our purpose was to find out which attentional modulations of the ERPs (e.g., N1, P1, RN and P300) are consistently associated with the vase–face illusion after face or vase cues. Based on previous work (e.g., Kornmeier and Bach, 2004, 2006; Pitts et al., 2007, 2008), we hypothesized that the perceptual reversal of ambiguous figure (the vase–face illusion) may not only depend on basic perceptual principles (low-level visual perception) but also be influenced by top-down control. That is to say, there might be some difference for these early and late ERP components (N1, P1, RN or P300) between the face and vase response to the vase–face figure.

2. Method

2.1. Subjects

Twelve junior undergraduates (6 women, 6 men) aged 19–24 years (mean age, 22.7 years) from Southwest University in China participated in the experiment as paid volunteers. All subjects gave written informed consent, were right-handed, had no current or past neurological or psychiatric illness, and had normal or corrected-to-normal vision.

2.2. Stimuli and procedure

The experiment included the classic vase–face ambiguous figure (see a in Fig. 1) and two modified disambiguous versions of the vase–face figure, including the face (see b in Fig. 1) and the vase (see c in Fig. 1) figures. Subjects viewed the stimulus which subtended 5.6° (height) × 7.4° (width) of visual angle, presented in the center of a 17-inch screen 80 cm from the subject at eye level (fixation at centre) in a quiet, well lit room.

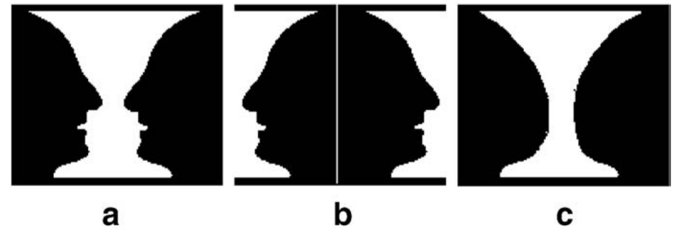

Fig. 1. Stimuli used in our experiment: the face–vase ambiguous figure (a), the face figure (b) and the vase figure (c).

The order of experimental procedure is as follows: the cue word (face or vase) appeared for 800 ms, the fixation point appeared for 300 ms after a 200–400 ms interval, then the stimulus (the vase–face, the face and the vase figures, randomized) appeared for 1500 ms, and an empty screen appeared for 1000 ms. Subjects were instructed to rest their right index and right middle finger on the 1 and 2 on the keyboard. When the cue word is face (or vase), the subjects were required to make a "yes" (press 1 key) response if there was indeed a face (vase) in the figure, while "no" (press 2 key) response if not. The experiment was divided into a practice phase and a test phase. When the participant was familiar with the procedure of the experiment, the practice phase was ended. The formal test consisted of six blocks, and every block had 60 judgement trials (10 trials for each condition, randomized). Subjects were instructed to avoid blinking and eye movement of any sort and to keep their eyes fixated on the monitor rather than looking down at their fingers during task performance. They were able to rest after finishing each block.

2.3. Electrophysiological recording and analysis

Brain electrical activity was recorded from 64 scalp sites using tin electrodes mounted in an elastic cap (Brain Product), with the reference on the left and right mastoids. The vertical electro-oculogram (EOG) was recorded with electrodes placed above and below the left eye. All interelectrode impedance was maintained below 5 kΩ. The EEG and EOG were amplified using a 0.05–80 Hz bandpass and continuously sampled at 500 Hz/channel for off-line analysis. Eye movement artifacts (blinks and eye movements) were rejected offline and 16 Hz low pass filter was used. Trials with EOG artifacts (mean EOG voltage exceeding ± 100 μV) and those contaminated with artifacts due to amplifier clipping, bursts of electromyographic activity, or peak-to-peak deflection exceeding ± 100 μV were excluded from averaging.

According to subjects' behavioral data, EEG of "face response" to the vase–face ambiguous figure after the face cue (face–face response or Face trials), "vase response" to the vase–face ambiguous figure after the vase cue (vase–vase response or Vase trials) was separately averaged. The averaged epoch for ERP was 600 ms including a 100-ms pre-stimuli baseline. On the basis of the ERPs grand averaged waveforms and topographical maps (see Figs. 2 and 3), the following 17 electrode points were chosen for statistical analysis: FPz, Fz, FCz, Cz, AF7, AF8, F1, F2, FC1and FC2 (10 sites for anterior); CPz, Pz, POz, CP1, CP2, P1 and P2 (7 sites for posterior). Latencies and amplitudes (baseline to peak) of the anterior N100, P160 and N320 were measured separately in the 80–120 ms, 140–200 ms and 260–350 ms time windows, respectively, and the posterior P100 and N160 were measured separately in the 80–120 ms and 140–200 ms time windows, respectively. Mean amplitudes in the time window of 350–450 ms was analyzed using two-way repeated-measures analyses of variance (ANOVA). The ANOVA factors were response type (face–face response, vase–vase response) and electrode site. For all analyses, p-values of all main and interaction effects were corrected using the Greenhouse–Geisser method for repeated-measures effects.