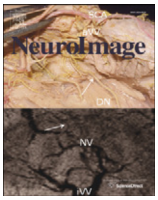
NeuroImage 81 (2013) 81–95

Contents lists available at SciVerse ScienceDirect

NeuroImage

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



Multiple synergistic effects of emotion and memory on proactive processes leading to scene recognition

Antonio Schettino [a,b,*], Tom Loeys [c], Gilles Pourtois [a]

[a] Department of Experimental-Clinical and Health Psychology, Ghent University, Ghent, Belgium
[b] Institute of Psychology I, University of Leipzig, Leipzig, Germany
[c] Department of Data Analysis, Ghent University, Ghent, Belgium

A R T I C L E I N F O

Article history:
Accepted 27 April 2013
Available online 9 May 2013

Keywords:
Scene recognition
Emotion
High density ERP
sLORETA
Parahippocampal gyrus
Fusiform gyrus
Dorsal anterior cingulate cortex
Insula

A B S T R A C T

Visual scene recognition is a proactive process through which contextual cues and top-down expectations facilitate the extraction of invariant features. Whether the emotional content of the scenes exerts a reliable influence on these processes or not, however, remains an open question. Here, topographic ERP mapping analysis and a distributed source localization method were used to characterize the electrophysiological correlates of proactive processes leading to scene recognition, as well as the potential modulation of these processes by memory and emotion. On each trial, the content of a complex neutral or emotional scene was progressively revealed, and participants were asked to decide whether this scene had previously been encountered or not (delayed match-to-sample task). Behavioral results showed earlier recognition for old compared to new scenes, as well as delayed recognition for emotional vs. neutral scenes. Electrophysiological results revealed that, ~400 ms following stimulus onset, activity in ventral object-selective regions increased linearly as a function of accumulation of perceptual evidence prior to recognition of old scenes. The emotional content of the scenes had an early influence in these areas. By comparison, at the same latency, the processing of new scenes was mostly achieved by dorsal and medial frontal brain areas, including the anterior cingulate cortex and the insula. In the latter region, emotion biased recognition at later stages, likely corresponding to decision making processes. These findings suggest that emotion can operate at distinct and multiple levels during proactive processes leading to scene recognition, depending on the extent of prior encounter with these scenes.

Introduction

Visual scene recognition results from dynamic and reciprocal interactions between bottom-up sensory processing and top-down modulatory influences, including selective attention (Treisman and Kanwisher, 1998), contextual information (Oliva and Torralba, 2007), and prior expectations (Summerfield and Egner, 2009). Interestingly, it has been proposed that the differential sensory processing of low (LSF) and high (HSF) spatial frequency information could underlie reciprocal interaction effects between bottom-up sensory processing and top-down expectations (Bar, 2003, 2004, 2007). In this framework, the rapid extraction of LSF information via dedicated magnocellular pathways may serve to quickly generate coarse predictions regarding the most probable content of the visual scene. This process would later be assisted and refined by the extraction of HSF information.

Whether these dynamic perceptual processes are differentially engaged depending on the (perceived) emotional content of the scenes, however, has received little attention so far. Mounting evidence shows that emotion exerts strong biases on visual perception (Öhman et al., 2001; Pourtois et al., 2013; Todd et al., 2012; Vuilleumier, 2005), including at early stages of recognition (Damaraju et al., 2009; Halgren et al., 2000; Pourtois et al., 2004). Thus, emotion is no longer seen as a byproduct of perception but, instead, as a core determinant of it (Pourtois et al., 2013). Accordingly, proactive processes during scene recognition are unlikely to be immune to emotion, defined here as the acquired valence and arousal values of the stimulus (Barrett et al., 2007; Lang, 1995; Russell, 1980). More specifically, the speed and extent of iterative processes between bottom-up processing and top-down expectations during scene recognition are probably subject to variations depending on the rapidly extracted emotional meaning of the stimulus. Consistent with this assumption, we recently reported evidence for the modulation of scene recognition processes by emotion (Schettino et al., 2011). In this study, participants were presented with series of pictures whose content was progressively revealed by increasing, in up to six sequential steps, the amount of LSF and HSF information. This *progressive unfolding task* was used to mimic a "coarse-to-fine" decomposition of the retinal input, presumably fostering the online generation of guesses regarding the most likely identity of the stimulus (Bar, 2003; Bullier, 2001; Hegdé, 2008).

---

* Corresponding author at: Institute of Psychology I, University of Leipzig, Seeburgstraße 14-20, 04103 Leipzig, Germany.
E-mail address: antonio.schettino@uni-leipzig.de (A. Schettino).

1053-8119/$ – see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.neuroimage.2013.04.115

Participants were asked to discriminate the content of the stimulus progressively revealed by performing an orthogonal animacy judgment task. Crucially, the scenes used in this experiment were neutral, pleasant, or unpleasant. Behavioral results showed a delayed recognition for emotional compared to neutral scenes. At the electrophysiological level, we found reliable ERP differences starting at approximately 280 ms after stimulus onset, depending on the amount of perceptual evidence accumulated. Estimated activity in the posterior cingulate cortex (PCC) and the parahippocampal gyrus (PHG) showed a response profile characterized by a linear and monotonic accumulation of perceptual evidence. Conversely, categorical recognition effects were evidenced in medial frontal regions, including the dorsal anterior cingulate cortex (dACC). However, none of these brain effects, albeit consistent with previous neuroimaging results (Carlson et al., 2006; James et al., 2000; Ploran et al., 2007, 2011), was found to be influenced by the emotional content of the scenes.

This lack of modulatory effect of emotion on these brain mechanisms was imputed to two main factors. First, it was hypothesized that the elected procedure did not effectively promote the generation of unambiguous predictions regarding the likely identity of the scenes progressively revealed, since participants were mainly presented with visually impoverished versions of pictures that they had never encountered before. Second, the use of an orthogonal animacy judgment task might have strongly reduced the salience of the emotional features of the scenes. Several studies showed that (early and automatic) affective stimulus processing is substantially reduced when concurrent, non-affective semantic stimulus dimensions become task-relevant (Everaert et al., 2011; Pessoa, 2008; Pessoa et al., 2002; Spruyt et al., 2007). To overcome these limitations, we devised a modified version of the progressive unfolding task (Schettino et al., 2012). In this latter study, to promote the generation and use of a restricted number of predictions regarding the identity of the scenes progressively revealed, we introduced a standard delayed match-to-sample manipulation (Courtney et al., 1997; Goldman-Rakic, 1990; Henson et al., 2005; Rugg et al., 1998). On each trial, participants were first asked to encode a complex colorful scene having either a neutral or emotional meaning. After a constant time interval, the content of either the same or a novel scene was progressively revealed (similarly to Schettino et al., 2011), and participants were asked to perform a delayed match-to-sample task. In addition, we occasionally asked participants to rate the emotional content of the scenes (besides the delayed match-to-sample task), thereby ensuring that the emotion features were directly attended throughout the whole experimental session. Behavioral results revealed earlier recognition for old compared to new scenes in all emotion conditions, providing evidence that participants used restricted (short-term) memory representations to perform the task. Notably, delayed recognition was observed for emotional relative to neutral scenes, replicating our previous behavioral results (Schettino et al., 2011).

The present paper reports the electrophysiological results of the aforementioned behavioral study (Schettino et al., 2012). Our main objective was to compare, using advanced EEG methods, the processing of previously encountered ("old") vs. "new" scenes, in order to evaluate whether any influence of emotion on brain mechanisms underlying scene recognition was depending on the (short-term) memory status of these complex visual stimuli. To verify that our short-term memory manipulation was effective, we analyzed a distinctive ERP signature of recollection processes, a late positive ERP waveform typically recorded from centro-parietal sites starting at approximately 500 ms post-stimulus onset (Curran and Cleary, 2003; Curran and Doyle, 2011; Rugg and Curran, 2007; Voss and Paller, 2006). Based on previous literature, we expected to observe more positive ERPs after correctly recognized old pictures compared to correctly classified new pictures, the so-called 'old/new effect' (Rugg and Curran, 2007; Voss and Paller, 2008; Wilding et al., 1995). Topographic analysis and complementary source localization methods were subsequently employed to investigate whether dissociable response profiles could be evidenced as a function of memory status and emotional valence of the scenes. Similarly to our previous study (Schettino et al., 2011), reliable topographical differences were expected to be observed in a time window following the low-level visual discrimination of the stimulus (i.e., after the N1 and P2 ERP components). In addition, we used a standard distributed source localization method to estimate the configuration of the neural generators of the main topographical maps identified by the preceding analysis. We assessed whether differential activation profiles in specific ventral brain areas as opposed to dorsal/medial frontal regions might be found, reflecting accumulation of evidence and decision-related processes, respectively.

Methods

Participants

Twenty-two undergraduate psychology students (all women, mean age 21 years, range 18–26) participated in the study, which was approved by the local Ethics Committee. All participants were native Dutch speaking, right-handed, had normal or corrected-to-normal vision, with no history of neurological or psychiatric disorders. The data of one participant could not be saved properly because of technical problems, whereas the EEG data of three other participants were discarded due to excessive artifacts and a low signal-to-noise ratio. Thus, the final sample consisted of 18 participants. All volunteers gave informed written consent prior to their participation, and were compensated 30 €.

Stimuli

The visual stimuli were selected from the International Affective Picture System (IAPS; Lang et al., 2008). Three-hundred and sixty pictures were pre-selected, equally divided into three emotion categories according to their pre-defined valence scores: *neutral*, *unpleasant*, and *pleasant* (see Table 1 in Schettino et al., 2012). Of note, these pictures were selected on the basis of mean valence and arousal ratings reported by female responders (Lang et al., 2008), because only women eventually participated in the experiment (see above). Similar to Schettino et al. (2011), we selected scenes that were neither highly pleasant (i.e., erotic situations) nor highly unpleasant (i.e., mutilations), because these categories were associated with specific emotion reactions in previous research (Schupp et al., 2006, 2007). We also included 16 additional neutral pictures that were only used during the practice session (not included in the subsequent statistical analyses). Furthermore, 36 supplementary neutral scenes were scrambled (i.e., each scene was divided into arbitrary grids of 255 × 255 pixels, whose locations were randomly shuffled 10 times) and, as a result, their content was made meaningless. Thus, in total, participants were presented with 412 IAPS scenes.[1]

The selected IAPS pictures were resized to 922 × 691 pixels (90% of the original 1024 × 768 pixel size), and they were presented at the beginning of each trial (see details below). For the progressive unfolding procedure, a separate pre-processing was carried out on the original IAPS scenes (1024 × 768 pixels) (Schettino et al., 2011). First, a grayscale conversion was applied using the algorithm included in Adobe Photoshop 6.0 (Adobe Systems, Mountain View, CA). Afterwards, six bandpass spatial frequency filters were applied on each and every picture using the fast Fourier transform (FFT) included in ImageJ v1.44 (http://rsb.info.nih.gov/ij/; for a similar procedure, see Delplanque et al., 2007). As a result, six distinct levels of filtering were obtained for every IAPS scene, each containing a different amount of low and high spatial frequency information.[2] All these modified pictures were finally resized to 768 × 576 pixels (75% of

---

[1] See Schettino et al. (2012) for a complete list of the IAPS pictures selected for the experiment.

[2] *Image1*: 256–1024 pixels/cycles; *Image2*: 128–1024 pixels/cycles; *Image3*: 64–1024 pixels/cycles; *Image4*: 32–1024 pixels/cycles; *Image5*: 16–1024 pixels/cycles; *Image6*: 0–1024 pixels/cycles. The numbers in the pair denote the pixels between which the cycles were kept after FFT (Delplanque et al., 2007).