

Running head: BRAIN DYNAMICS OF VISUAL OBJECT RECOGNITION

**Brain Dynamics of Upstream Perceptual Processes Leading to Visual Object
Recognition: A High Density ERP Topographic Mapping Study**

Antonio Schettino¹, Tom Loeys², Sylvain Delplanque³ & Gilles Pourtois¹

(1) Department of Experimental-Clinical and Health Psychology, Ghent University; (2)
Department of Data Analysis, Ghent University; (3) Swiss Center for Affective Sciences,
University of Geneva

Corresponding author:

Gilles Pourtois

Department of Experimental-Clinical and Health Psychology

Ghent University

Henri Dunantlaan 2

9000 Ghent

Belgium

Phone: +32 9 264 9144

Email: gilles.pourtois@ugent.be

Abstract

Recent studies suggest that visual object recognition is a proactive process through which perceptual evidence accumulates over time before a decision can be made about the object. However, the exact electrophysiological correlates and time-course of this complex process remain unclear. In addition, the potential influence of emotion on this process has not been investigated yet. We recorded high density EEG in healthy adult participants performing a novel perceptual recognition task. For each trial, an initial blurred visual scene was first shown, before the actual content of the stimulus was gradually revealed by progressively adding diagnostic high spatial frequency information. Participants were asked to stop this stimulus sequence as soon as they could correctly perform an animacy judgment task. Behavioral results showed that participants reliably gathered perceptual evidence before recognition. Furthermore, prolonged exploration times were observed for pleasant, relative to either neutral or unpleasant scenes. ERP results showed distinct effects starting 280 ms post-stimulus onset in distant brain regions during stimulus processing, mainly characterized by: (i) a monotonic accumulation of evidence, involving regions of the posterior cingulate cortex/parahippocampal gyrus, and (ii) true categorical recognition effects in medial frontal regions, including the dorsal anterior cingulate cortex. These findings provide evidence for the early involvement, following stimulus onset, of non-overlapping brain networks during proactive processes eventually leading to visual object recognition.

Keywords: visual object recognition, accumulation of evidence, dorsal ACC, parahippocampal gyrus, posterior cingulate cortex

Brain Dynamics of Upstream Perceptual Processes Leading to Visual Object Recognition: A
High Density ERP Topographic Mapping Study

Introduction

Visual object recognition is a fast, accurate and effortless process. Despite swift or dramatic variations in the retinal input (i.e. due to modifications in orientation, size, appearance, viewpoint, or context), human's ability to readily detect and recognize a multitude of visual objects in the environment is hardly challenged (Bar, 2003; Biederman & Bar, 1999). Many studies and models have emphasized the sequential property of visual computations leading to perceptual decision making, from the analysis of sensory information to the selection of the behavioral outcome that best maximizes the expected utility (Biederman, 1987; Marr, 1982; Opris & Bruce, 2005). According to the dominant framework, the visual system evaluates in a probabilistic fashion the available information about various features of the input image, thereby making inferences about its content and preparing possible courses of action (Hegd , 2008). Classical neurophysiological models postulate that fundamental visual features of the input image are initially processed in lower-level cortical areas of the occipital and infero-temporal cortex, after which they are used to generate an abstract visual representation of the object. Recognition (and subsequent motor execution of the congruent response) is achieved when this representation is successfully matched with templates stored in memory (Biederman, 1987; Palmer, 1999; Ranganath & Rainer, 2003; Riesenhuber & Poggio, 1999; Ungerleider & Mishkin, 1982).

The predominant role of bottom-up perceptual processes leading to recognition is clearly stressed in these hierarchical models. However, top-down modulatory effects may also influence visual object recognition mechanisms. They include selective attention (Blair, Watson, Walshe,

& Maj, 2009; Treisman & Kanwisher, 1998), task relevance (Egner & Hirsch, 2005), prior probability of encounter (Summerfield & Egner, 2009), working memory (Ranganath, Cohen, Dam, & D'Esposito, 2004), contextual information (Bar, 2004; Oliva & Torralba, 2007), as well as the monitoring of the decision's outcome (Ridderinkhof, Ullsperger, Crone, & Nieuwenhuis, 2004; Ullsperger, Volz, & von Cramon, 2004). Hence, the rapid bottom-up processes leading to visual object recognition could be assisted by the online activation of abstract information -- stored in higher-level brain regions -- primarily recruited to speed up the concurrent ongoing processing in lower-level visual areas, with the aim to limit the number of computations necessary to eventually identify an object (Bar et al., 2006; Ganis, Schendan, & Kosslyn, 2007).

Recent theoretical accounts emphasized the role of expectations in visual recognition (Bar, 2009; Kersten, Mamassian, & Yuille, 2004; Summerfield & Egner, 2009; Yuille & Kersten, 2006). In these models, it is hypothesized that the active use of prior information about the most probable visual percept in the forthcoming sensory environment is at work in order to guide the rapid acquisition of diagnostic visual information (i.e. invariant and expected aspects in the environment do not need to be processed thoroughly), as well as to facilitate the interpretation of ambiguous stimuli. Predictive coding models of visual recognition (Friston, 2005; Friston & Kiebel, 2009; Grossberg, 2009; Rao & Ballard, 1999; Serences, 2008) argue that feed-forward information coming from early visual areas is compared, at each stage of the visual processing hierarchy, to top-down expectations whose aim is to predict the responses at the next lower level, primarily through recurrent or feedback connections (see also Di Lollo, Enns, & Rensink, 2000; Enns & Lleras, 2008). If this comparison results in an erroneous output, an error signal is sent back to the higher level via feedforward connections. This error signal is then used to correct the estimate of the input signal at each level (Rao & Ballard, 1999; Serences, 2008). Expected and

observed information are iteratively adjusted until the visual system is able to settle on a single perceptual interpretation of the sensory input (Summerfield & Egnér, 2009). A plausible neural mechanism underlying the triggering of this top-down facilitation in object recognition has recently been proposed by Moshe Bar (Bar, 2003, 2004, 2007, 2009). According to this model, a partially analyzed version of the input image, mainly comprised of low spatial frequency (LSF) information, is projected rapidly from early visual cortex directly to orbitofrontal (OFC) and parahippocampal (PHC) areas, possibly via a fast dorsal magnocellular route. In the PHC, this blurred image activates the most probable (experience-driven) guesses about the context frame that needs to be triggered. This contextual information is projected to the infero-temporal cortex, where a set of visual associations corresponding to the relevant context is activated. In parallel, the visual information conveyed by the same scene -- here the allocation of attention is on the target object -- would be sufficient to rapidly activate the most likely interpretations of the input image in the OFC. The integration of the representations of the specific context and the candidate interpretations of the target object would in turn result in the reliable selection of a single identity, which can be further refined with specific detail gradually conveyed by high spatial frequency (HSF) information (Bar, 2004). Consistent with this framework, Peyrin et al. (2010) combined fMRI and ERPs to explore the prioritization of LSF in the processing of visual input, and found that higher order areas in frontal and temporo-parietal regions responded more to LSF stimuli when presented first, whereas occipital visual cortex responded more to HSF stimuli when presented after LSF stimuli (Bar, 2003; Bullier, 2001; Hegdé, 2008). Hence, the use of different imaging methods confirmed a “coarse-to-fine” processing of visual input (Hegdé, 2008; Lomber, 2002; Navon, 1977; Sanocki, 1993; Schyns & Oliva, 1994): the quick processing of

LSF in higher order frontal or temporal brain regions could directly influence scene recognition by biasing perceptual processes in object-selective visual areas.

Expectations may lower the threshold that needs to be overcome to make a decision in favor of one option, similar to what accumulator models of decision making propose (Gold & Shadlen, 2007; Ratcliff & McKoon, 2008). Such models have proven highly effective in describing performance in recognition memory (Ratcliff, 1978), economic decisions (Sanfey, Loewenstein, McClure, & Cohen, 2006), semantic (Ratcliff & McKoon, 1982) and lexical (Ratcliff, Thapar, Gomez, & McKoon, 2004) tasks, as well as sensory discrimination (Gold & Shadlen, 2007). According to these models, incoming information is accumulated over time, evaluated and assigned to a response option: the actual decision is made when evidence in favor of one of the two responses exceeds a threshold. This framework has been successfully applied in studies on nonhuman primates (Hanes & Schall, 1996; Kim & Shadlen, 1999; Shadlen & Newsome, 2001; Tanaka, 1996), eventually providing a useful mean to link neurophysiology with behavior. Important hints on the brain regions underlying proactive perceptual processes leading to visual object recognition in humans have also been obtained more recently, using mainly fMRI and gradual or progressive stimulus revelation tasks (Carlson, Grol, & Verstraten, 2006; Heekeren, Marrett, Bandettini, & Ungerleider, 2004; Heekeren, Marrett, & Ungerleider, 2008; James, Humphrey, Gati, Menon, & Goodale, 2000; Ploran et al., 2007; Wheeler, Petersen, Nelson, Ploran, & Velanova, 2008). Among these studies, Ploran, et al. (2007) used an elegant procedure enabling a gradual revelation of masked stimuli over eight discrete and consecutive steps. Participants had to press a button as soon as they could identify the pictures' content with a reasonable degree of confidence. Interestingly, Ploran, et al. (2007) found that different brain areas showed distinct effects during this progressive revelation task. Three different patterns of

brain activation were identified. For several posterior occipital regions, activity increased monotonically as a function of the amount of visual information entering the visual system (sensory processors). By contrast, a gradual increase in activity with a distinctive peak at the time of recognition was found in inferior temporal, frontal and parietal regions (including the bilateral fusiform gyrus and the DLPFC), consistent with an accumulation of evidence process which may be necessary to recognize the identity of the target object (accumulators). Finally, in many regions of medial frontal cortex (including the dorsal region of the anterior cingulate cortex, dACC, and the anterior insula), activity remained close to baseline until the moment of actual recognition, suggesting their involvement in decision-related processes that accompany overt visual object recognition. These latter fMRI results are important, as they inform about upstream brain mechanisms leading to visual object recognition and their different functions, with non-overlapping areas involved in accumulation of evidence vs. moment of recognition (Ploran et al., 2007). However, since these results were obtained using fMRI techniques, the electrophysiological correlates and actual time-course of these upstream perceptual or decision-related effects during visual object recognition remain unclear.

The goal of this study was to use a novel stimulus revelation task and explore, using high density EEG in healthy adult participants, the precise electrophysiological correlates of upstream processes leading to visual object recognition. More specifically, we sought to investigate whether different ERP response profiles could be evidenced, consistent with the assumption of accumulation of evidence prior to visual object recognition subserved by posterior occipital or temporal brain regions, in comparison with medial frontal areas which might show categorical moment-of-recognition effects (see Ploran, et al., 2007). For this purpose, we designed a new task enabling a progressive revelation of the stimulus content, while neural events prior to actual

recognition were investigated on a trial-by-trial basis using the millisecond time resolution provided by EEG. Based on the psychological models and neuroscience evidence reviewed here above, we hypothesized that upstream perceptual effects leading to visual object recognition could be twofold (Gold & Shadlen, 2007; Ratcliff & McKoon, 2008). First, we surmised ERP effects reflecting the accumulation of perceptual evidence over time, with the main neural generators being localized in posterior brain regions, including the occipital and temporal cortices (Ploran et al., 2007). Consistent with this view, a linear relationship between the electrophysiological signal and the amount of visual input was expected. In contrast, we predicted that medial frontal brain regions (e.g. dorsal ACC, see Ploran, et al., 2007) might also provide an important mechanism at stake during visual object recognition, showing moment-of-recognition effects characterized by a non-linear and abrupt change in the amplitude of the ERP signal occurring close to actual recognition. As a secondary question, we also investigated whether the emotional content of the stimulus/scene might influence these upstream perceptual processes leading to recognition, consistent with previous imaging studies showing reciprocal interaction effects between regions involved in the processing of emotions (including the amygdala) and more posterior regions implicated in visual object recognition (Sabatinelli, Lang, Bradley, Costa, & Keil, 2009; Vuilleumier, 2005). Hence, we presented participants with standard neutral and emotional scenes (whose content was progressively revealed) and tested for a differential neural effect likely triggered by the emotional content prior to actual recognition. Whereas previous studies have found reliable modulation of early and late ERP components following stimulus onset as a function of the emotional content of the stimulus (either for valence or arousal dimensions, see Schupp, et al., 2000; Schupp, Junghöfer, Weike & Hamm, 2003; Carretié, Hinojosa, Martin-Loeches, Mercado, & Tapia, 2004; D'Hondt, et al., 2010;

Delplanque, Lavoie, Hot, Silvert, & Sequeira, 2004; Oloffson, Nordin, Sequeira, & Polich, 2008), to our knowledge, no ERP study has systematically explored whether upstream or anticipatory perceptual effects during visual object recognition may be influenced by the rapid decoding of the emotional content of the input stimulus, likely based on the selective processing of low-spatial frequency/impooverished visual cues. Accordingly, the goal of our study was also to verify whether the rapid processing of the emotional content of the scene (presumably based on LSF cues; see Vuilleumier, Armony, Driver, & Dolan, 2003; Pourtois, Dan, Grandjean, Sander, & Vuilleumier, 2005; Bar, 2003, 2004) could alter upstream brain processes underlying proactive mechanisms of visual object recognition.

Methods

Participants

Nineteen psychology students (14 women, mean age 21 years, range 17-33) were recruited to freely participate in the study, which was approved by the local university ethical committee. All participants were native Dutch speaking, right-handed, had normal or corrected-to-normal vision, with no history of neurological or psychiatric disorders. All volunteers gave informed written consent prior to their participation, and were paid 20€.

Stimuli

Two hundred and thirty-four pictures were selected from the International Affective Picture System (IAPS; Lang, Bradley & Cuthbert, 2005), a standardized database containing emotionally-evocative pictures that depict objects and scenes across a wide range of categories and situations. The IAPS manual provides normative values for the basic dimensions of emotion -- including arousal and valence -- as rated by the Self-Assessment Manikin (SAM) on a scale from 1 to 9 (Bradley & Lang, 1994). Since there are gender differences in both valence and

arousal ratings, we selected two sets of pictures in order to balance the arousal levels of the emotional pictures across male vs. female participants. Each set consisted of 138 pictures. Among these, 42 were shared between male and female participants. The pictures were divided into three emotion categories, according to their pre-defined valence scores: neutral, unpleasant and pleasant (see Table 1). For each category (N = 46), half of the pictures contained living objects (i.e. human beings or animals) while the other half did not (i.e. landscapes or artifacts)¹. We explicitly selected pictures that were neither highly pleasant (i.e. erotic scenes) nor highly unpleasant (i.e. mutilations), because such pictures could lead to specific emotion reactions which may be different between male and female participants (Lithari et al., 2010; Proverbio, Adorni, Zani, & Trestianu, 2009; Sabatinelli, Flaisch, Bradley, Fitzsimmons, & Lang, 2004). We also selected ten additional neutral pictures that were used during the practice session (see footnote) and were not included in the subsequent statistical analyses. Furthermore, 18 pictures were scrambled and their content made meaningless. These scrambled pictures were eventually used as “catch” trials to ensure that participants correctly attended to the content of the pictures before taking a decision.

The pre-selected original IAPS pictures (1024 x 768 pixels, corresponding to 20° x 15° of visual angle at a distance of 75 cm) were first converted to grayscale. Each picture was then bandpass filtered (using ImageJ v1.44 software, <http://rsb.info.nih.gov/ij/>) according to the spatial frequency bands put forward by Delplanque, N'diaye, Scherer, & Grandjean (2007) (see Table 2). This procedure resulted in six different levels of filtering for each and every picture. Each of these levels reliably differed from one another depending on the actual content of low and high spatial frequency information (see Figure 1A).

Procedure

Participants were tested individually in a small, dimly lit room, and seated at 75cm in front of a 19" CRT computer screen (refresh rate 100 Hz). They were first asked to complete a practice/familiarization block containing 10 trials. Then, they started the experimental session, which was divided into four blocks (separated by a one-minute pause), each containing 67-68 trials. Each trial had the same underlying structure (see Figure 1A). It began with a fixation cross displayed for 250ms in the center of the screen. Then the first (blurred) image level of a given picture was presented for 500ms, followed by a 250ms blank screen. Next, the second image level of the same picture (containing more HSF information) was immediately presented for 500 ms, plus 250 ms blank screen, and the same procedure was repeated until the presentation of the sixth image level (i.e. intact/unfiltered picture). Hence, this procedure resulted in a progressive and predictive revelation of the image content by adding, in a stepwise fashion, high spatial frequency information to an initial blurred and meaningless picture. The inter-trial interval (ITI) was set to 1000ms. Participants were instructed to provide two consecutive responses. First, they were asked to press the spacebar key (on a standard AZERTY keyboard) with their dominant right hand as soon as they felt they could decide, with sufficient confidence, whether the scene contained a living object or not (Response1). Pressing the spacebar key immediately interrupted the presentation of the stimuli. 500 ms after pressing the spacebar, participants were required to validate their choice and to perform a two-alternative forced choice task. They were asked to press the "L" key of the keyboard if the scene contained a living object, or alternatively the "N" key if it did not contain any living object (Response2). This dual registration procedure enabled us to timely separate early recognition effects (Response1) from the overt discrimination of the scene (Response2). Importantly, the actual discrimination (Response2) was required to distinguish correct from incorrect early key presses (when looking retrospectively at ERP data

recorded around the onset of Response1, see here below). Hence, ERP analyses were primarily focused on neural events taking place prior to Response1, when the actual visual discriminations (based on Response2) turned out to be accurate and errors were removed from the analyses. Because Response1 always required a simple key press, this procedure minimized the potential contamination of ERP data by the activation of competing responses (living vs. non-living). Participants were instructed to find a good balance between high accuracy and the execution of speeded responses. Note that they were not encouraged to respond before the presentation of the last/sixth (unfiltered) picture. For catch trials (i.e. scrambled pictures), a non-living response (Response2) was expected.

Since we were primarily interested in ERP effects which might foreshadow the actual detection and recognition of a scene, we had, as a prerequisite, to include enough trials per condition (neutral, unpleasant and pleasant) to be able to eventually compute reliable ERP waveforms per condition and for each image level separately. However, we only had a limited set of pictures which were balanced with regard to the living vs. non-living attribute. For this reason, unbeknown to participants, each picture was presented twice during the experimental session. There was a random and unpredictable time lag (i.e. 4-15 intervening images, $\underline{M} = 10$, $\underline{SD} = 3$) between first presentations and repetitions. As expected, behavioral results showed significant earlier scene recognition for repetitions, relative to first presentations, and this priming effect was the same for neutral, pleasant and unpleasant scenes (all $ps < .05$). However, separate analyses of the ERP data for these two conditions (first presentations vs. repetitions) did not reveal any significant difference. Accordingly, they were collapsed in the statistical analyses to increase the signal-to-noise ratio.

Stimulus presentation and behavioral response recordings were controlled using E-Prime software (V2.0., <http://www.pstnet.com/products/e-prime/>).

Questionnaires

At the end of the experimental session, participants were asked to fill out three different questionnaires, in order to assess whether specific affective or personality traits might be related to task performance and/or our ERP effects. Trait anxiety levels were measured by means of the Dutch version of the State-Trait Anxiety Inventory, trait characteristics (Van der Ploeg, Defares, & Spielberger, 1979). Participants also completed a recent Dutch version of the Need For Closure Scale (Roets & Van Hiel, 2007), which gave an estimate of the participant's tendency or need to obtain any answer, as opposed to tolerating ambiguity. Finally, we also administered the Need For Affect Scale (Maio & Esses, 2001) in order to obtain an independent measure of the general motivation of participants to either approach or avoid situations that are emotion-inducing. However, we found no significant correlation between the scores obtained for each of the three questionnaires and either the behavioral or ERP results obtained during this progressive stimulus revelation task, presumably because of the low standard deviation of the scores obtained for each questionnaire (see Table 3).

EEG data acquisition and pre-processing

Electroencephalographic (EEG) activity was continuously recorded using a BIOSEMI Active-Two system (BioSemi, Inc., Netherlands; <http://www.biosemi.com>) by means of 128 active electrodes fitted into a stretching cap and following the BioSemi ABCD positioning system (i.e. electrode positions are radially equidistant from CZ; <http://www.biosemi.com/headcap.htm>). Two electrodes, the common mode sense (CMS) active electrode and the driven right leg (DRL) passive electrode, were used as reference and ground electrodes, respectively

(http://www.biosemi/faq/cms_and_drl.htm). Vertical electro-oculograms (EOG) were monitored using two additional electrodes placed in the inferior and superior areas of the left orbit. EEG and EOG recordings were sampled at 512 Hz.

ERPs of interest were computed offline using Brain Vision Analyzer 2.0 (Brain Products™ GmbH, Munich, Germany; http://www.brainproducts.com/analyzer2_release.php). First, a common average reference was applied. Next, -100/+750 ms epochs were created around the onset of the visual stimulus. Afterwards, all the segments were baseline corrected using the pre-stimulus interval (100ms), before ocular correction was performed (Gratton, Coles, & Donchin, 1983). Artifact rejection was then carried out (mean amplitude of $\pm 76.32\mu\text{V}$ scale across participants) to eliminate segments contaminated by artifacts such as residual blinks or muscle activity. Approximately one-third of the data (27.9%) were discarded using these strict criteria. Finally, averaging of the segments was performed, separately for each condition, and a 1-30 Hz bandpass filter was applied to the individual averaged data. Since we were primarily interested in neural processes occurring before the actual recognition of the scene, we used the time of detection (Response1) as the initial reference point in our analyses, and looked at stimulus-locked ERP effects backwards. Only recognitions (Response1) which turned out to be accurate (based on Response2) were included in our ERP analyses (see behavioral results below). Following this procedure, four different epochs were computed for each individual scene: (1) -100/+750 ms around the stimulus onset that was recognized during this specific time window (“Recognition”); (2) -100/+750 ms around the stimulus onset which immediately preceded recognition (“One image before” recognition); (3) -100/+750 ms around the stimulus onset for the stimulus appearing two images before recognition (“Two images before” recognition); (4) -100/+750 ms around the stimulus onset for the stimulus appearing three images before

recognition (“Three images before” recognition). Epochs of 750 ms following stimulus were used because they encompassed the duration of the stimulus itself (500 ms), as well as the subsequent 250ms interval. Using this procedure, we could thus look at stimulus-locked ERP effects for image levels that preceded actual recognition, with a gradual distance relative to this event (up to three images before recognition). Note that this specific data analysis allowed us to look at brain processes consistently foreshadowing visual object recognition, bearing in mind that, across trials, the actual decision could be based on different accumulations of perceptual evidence (i.e. the proportion of correct Responses¹ turned out to be the largest at image levels three, four and five, see behavioral results below). Furthermore, we computed different ERP waveforms as a function of the emotional valence of the scenes. As a result, 12 individual averages were computed for each participant : (1) Neutral, Recognition (number of segments after pre-processing: $\underline{M} = 50$, $\underline{SD} = 9$); (2) Neutral, One image before ($\underline{M} = 58$, $\underline{SD} = 9$); (3) Neutral, Two images before ($\underline{M} = 58$, $\underline{SD} = 8$); (4) Neutral, Three images before ($\underline{M} = 43$, $\underline{SD} = 9$); (5) Unpleasant, Recognition ($\underline{M} = 44$, $\underline{SD} = 8$); (6) Unpleasant, One image before ($\underline{M} = 52$, $\underline{SD} = 7$); (7) Unpleasant, Two images before ($\underline{M} = 49$, $\underline{SD} = 7$); (8) Unpleasant, Three images before ($\underline{M} = 37$, $\underline{SD} = 10$); (9) Pleasant, Recognition ($\underline{M} = 47$, $\underline{SD} = 10$); (10) Pleasant, One image before ($\underline{M} = 56$, $\underline{SD} = 10$); (11) Pleasant, Two images before ($\underline{M} = 52$, $\underline{SD} = 9$); (12) Pleasant, Three images before ($\underline{M} = 47$, $\underline{SD} = 10$).

Analysis of behavioral data

Accuracy was expressed as percentage of correct responses. Since these image levels were not independent of each other (each image relies on the visual information conveyed by previous levels), cumulative percentages were calculated. Using this procedure, we obtained a psychometric curve showing the evolution of the recognition accuracy across the six image

levels for each condition separately. Repeated measures analyses of variance (ANOVAs) with two factors (emotion: three levels; image level: six levels) and post-hoc t-tests were first employed to assess differences in recognition accuracy between conditions. The level of significance for all the analyses was set to $p < 0.05$. In addition, to verify whether the psychometric curve was shifted as a function of the emotional content of the scene (pleasant, neutral or unpleasant), we also used a proportional odds model (Agresti, 2007), a regression model for ordinal dependent variables (recognition from image level 1,..., recognition from image level 6). This data analysis allows to model the cumulative probability up to and including recognition from image level k ($k = 1, 2, 5$). The derived odds ratio expresses how much the odds of recognition from image level k or earlier is increased (if larger than 1) or decreased (if smaller than 1) across different emotional contents, and thus provides a single number capturing the shift in psychometric curve. To account for dependencies of trials within the same subject, a multi-level version of the proportional odds model was used here.

Analyses of ERP data

In this study, reference-free topographic analyses were carried out to objectively characterize differences between conditions and image levels. The basic principles of this method have been described extensively elsewhere (see Lehmann & Skrandies, 1980; Michel, Seeck, & Landis, 1999; Michel, et al., 2001; Murray, Brunet, & Michel, 2008; Pourtois, Delplanque, Michel, & Vuilleumier, 2008). The added value of this method, relative to a more traditional ERP peak analysis (see Picton, et al., 2000), is that it enables to reveal global differences between experimental conditions without a priori selecting a few channels or time frames. All channels and time frames are used concurrently in the analysis. Topographic analyses were performed using CARTOOL software (version 3.43; <http://brainmapping.unige.ch/Cartool.htm>). For each

participant separately, we first interpolated noisy channels using a spherical splines transformation (Perrin, Pernier, Bertrand, & Echallier, 1989). Then, the dominant topographic scalp maps were identified in the grand average ERP data (Recognition, One image before, Two images before and Three images before) over a wide time-window spanning from 0 to 600ms after stimulus presentation. To this end, a specific spatiotemporal clustering algorithm, the “Atomize and Agglomerate Hierarchical Clustering” (AAHC; for a detailed description, see Murray, Brunet, & Michel, 2008), was used. This clustering method was developed on purpose to reduce complex EEG/ERP data sets. Following standard practice, the optimal number of dominant maps “explaining” the dataset was based on a cross-validation criterion (Pascual-Marqui, Michel, & Lehmann, 1995). The dominant scalp topographies (identified in the group-averaged data) were then fitted to the ERPs of each individual subject using spatial fitting procedures, to determine their representation across subjects and conditions. This procedure provided fine-grained quantitative values, such as the duration of a specific topographic map or its global explained variance (GEV, or goodness of fit), which are critical indices of the significance of a given topography, not available otherwise in a classical component analysis (Picton et al., 2000). GEV represents the sum of the explained variance weighted by the GFP (Global Field Power) at each moment in time. GEV was entered in repeated measures analyses of variance (ANOVAs) with image level and emotional content as within-subject factors. Paired t-tests (and a conservative Bonferroni correction) were used as post-hoc comparisons between conditions. When appropriate, degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity. As for the statistical analysis of the behavioral data, the level of significance for all these analyses was set to $p < 0.05$.

Source localization analyses

To estimate the likely neural sources underlying the electrical field configurations identified by the previous analysis, we used a specific distributed linear inverse solution, namely standardized low-resolution brain electromagnetic tomography (sLORETA: Pascual-Marqui, 2002). Mathematical validation of this distributed source localization technique has been recently demonstrated (Sekihara, Sahani, & Nagarajan, 2005). The head model for the inverse solution uses the electric potential lead field computed with a boundary element method applied to the MNI152 template (Fuchs, Kastner, Wagner, Hawes, & Ebersole, 2002). Scalp electrode coordinates on the MNI brain are derived from the international 5% system (Jurcak, Tsuzuki, & Dan, 2007). The source locations were therefore given as (\underline{x} , \underline{y} , \underline{z}) coordinates (\underline{x} from left to right; \underline{y} from posterior to anterior; \underline{z} from inferior to superior). The calculation of all reconstruction parameters was based on the computed common average reference. sLORETA units were scaled to amperes per square meter (A/m^2). Direct statistical comparisons between conditions were performed in this inverse solution space using paired t-test. The level of significance for all the analyses was set to $p < 0.05$.

Results

Accuracy

Participants were accurate in this task; mean error rate was 12.84% ($SD = 5.02$). A 3 (emotion) x 6 (image level) repeated measures ANOVA revealed a highly significant main effect of image level [$F(5, 108) = 36.09$, $p < .001$, $\eta_p^2 = .626$]. However, there was no evidence of either a significant main effect of emotion [$F(2, 36) = 0.24$, $p = .683$, $\eta_p^2 = .013$] or an emotion x image level interaction [$F(10, 180) = 1.33$, $p = .271$, $\eta_p^2 = .069$]. Post-hoc comparisons showed a significantly higher percentage of errors during the presentation of Image4 compared to the other image levels (all $ps < .001$), because the majority of responses actually occurred during the

presentation of this image level (see below). No significant difference was found between errors during Image1 and Image2 ($p > .05$). Furthermore, the percentage of errors was similar for Image1, Image2 and Image6 (all p s $> .05$). In addition, very few errors were committed with catch trials (Image1: $M = 0.00$, $SD = 0.00$; Image2: $M = 0.00$, $SD = 0.00$; Image3: $M = 0.44$, $SD = 1.91$; Image4: $M = 2.90$, $SD = 12.62$; Image5: $M = 5.70$, $SD = 15.73$; Image6: $M = 1.05$, $SD = 4.59$), confirming that participants reliably processed the content of the scene before making a response (living vs. non-living object).

Cumulative percentages of correct responses (i.e. Response1 only when Response2 was correct, see Method) are presented in Table 4. A 3 (emotion) x 6 (image level) repeated measures ANOVA performed on these values showed a significant main effect of emotion [$F(2, 36) = 39.62$, $p < .001$, $\eta_p^2 = .688$], a significant main effect of image level [$F(5, 90) = 637.37$, $p < .001$, $\eta_p^2 = .973$] and a significant emotion x image level interaction [$F(10, 180) = 21.80$, $p < .001$, $\eta_p^2 = .548$]. Post-hoc comparisons confirmed a progressive (although non-linear) gain in recognition from Image1 to Image6 (all p s $< .05$, with the exception of Image1 vs. Image2 for pleasant stimuli, $p > .05$), as shown by an S-shaped psychometric function (see Figure 1B). Hence, image levels three, four and five presumably provided sufficient diagnostic low and high spatial frequency information to perform the animacy judgment task with high accuracy. More generally, this S-shaped psychometric function confirmed that our progressive stimulus revelation task was successful, since participants did not respond randomly to the different filtered stimuli across trials but, instead, they consistently waited at least until Image3 before interrupting the stimulus sequence. Importantly, the analysis of the cumulative percentage of accuracy of catch trials revealed a similar outcome. A univariate ANOVA disclosed a significant main effect of image level [$F(5, 108) = 152.88$, $p < .001$, $\eta_p^2 = .876$]. Post-hoc comparisons

confirmed a progressive gain in recognition at each image level (all p s < .001, except Image1 vs. Image2, $p > .05$), lending additional support to the assumption that image levels three, four and five contained sufficient diagnostic spatial frequency information to perform the task with high accuracy. Altogether, these results are consistent with the assumption of non-linear accumulation of evidence prior to recognition (Biederman, 1987; Gold & Shadlen, 2007; Ratcliff & McKoon, 2008; Smith & Ratcliff, 2004).

Next, we assessed whether the emotional content of the scene may have had an influence on the actual recognition of these scenes. To this end, a mixed proportional odds model with emotion as fixed factor (neutral, unpleasant or pleasant), and subject as random effect was carried out (see Table 5). Interestingly, this analysis suggested a highly significant main effect of emotion ($p < .001$). Pairwise comparisons revealed a shift of the distribution as a function of the emotional content of the scene, indicated by reliably earlier animacy judgments when the picture contained a neutral, as opposed to either an unpleasant ($p < .001$) or pleasant ($p < .001$) content (Figure 1B). The psychometric function was also shifted leftwards (i.e. earlier recognition) for unpleasant relative to pleasant scenes ($p < .001$). Hence, these results showed that pleasant scenes were recognized significantly later than either unpleasant or neutral scenes.

Reaction Times

Finally, we calculated the mean reaction times (RTs) for the correct responses at each image level separately (Image1: \underline{M} = 0.00 ms, \underline{SD} = 0.00; Image2: \underline{M} = 133.21 ms, \underline{SD} = 217.89; Image3: \underline{M} = 476.43 ms, \underline{SD} = 104.22; Image4: \underline{M} = 362.00 ms, \underline{SD} = 54.59; Image5: \underline{M} = 288.32 ms, \underline{SD} = 40.29; Image6: \underline{M} = 194.18 ms, \underline{SD} = 105.94) and compared RTs at Image3, Image4 and Image5, for which the majority of behavioral responses were recorded. Pairwise comparisons revealed significant differences between Image3 and Image4 [$t(18) = 4.72$, $p <$

.001], Image3 and Image5 [$t(18) = 6.98, p < .001$] and Image4 and Image5 [$t(18) = 5.58, p < .001$], each time indicated by faster decisions for images levels containing more high spatial frequency information than levels providing more degraded visual information.

ERP results

Following standard practice, a spatiotemporal cluster analysis was applied on the four main ERP conditions (Recognition, One image before, Two images before and Three images before) during a broad temporal window following stimulus onset (0-600 ms post-stimulus). This analysis revealed that five distinct dominant field topographies explained 91.37% of the total variance (see Figure 2A and 2B). The two first dominant maps were common to all four conditions, with reliable topographic changes between conditions starting 280 ms post-stimulus onset. The first dominant map found in the spatial cluster analysis had a prolonged duration (0-216 ms post-stimulus onset) and shared several electrophysiological properties with the visual N1 component (Figure 2C) (Vogel & Luck, 2000). The fact that this ERP component, which is usually phasic, showed here a sustained effect may be explained by our specific task parameters, in which degraded stimuli are presented and a progressive accumulation of visual information is needed over a prolonged period of time. This N1 scalp map was next replaced by a visual P2 component (Carretié et al., 2004; Crowley & Colrain, 2004; Freunberger, Klimesch, Doppelmayr, & Holler, 2007; Luck & Hillyard, 1994) which, like the preceding N1 scalp map, was shared across the four image levels (Figure 2C). The spatiotemporal cluster analysis disclosed that the P2 scalp map had the highest variance during the 216-280 ms time interval post-stimulus onset. Most likely, these two early neural activities reflected the encoding and low-level visual discrimination of the stimulus. Following the offset of the P2 scalp map (280 ms post-stimulus onset), the cluster analysis unambiguously revealed reliable topographic changes

across the four image levels, which are necessarily indicative of changes in the configuration of the underlying intracranial generators (Lehmann & Skrandies, 1980; Michel et al., 1999; Murray et al., 2008). Whereas another broad occipito-parietal activity was generated following the P2 for image levels one, two and three before recognition, a distinctive occipital negative component (with a concurrent positive, frontal counterpart) was generated during the same latency for the image level corresponding to actual recognition (Figure 2C). Interestingly, this differential scalp topography was actually generated during the time period (280-360 ms post-stimulus, Recognition level) in which the first decision (Response1) was most likely made by the participant. However, since RTs were variable or jittered across trials and participants, as well as differed as a function of image level (see behavioral results above), it is unlikely that this distinctive scalp map was somehow related only to the preparation or execution of a motor response.

The next step was to verify whether the topographic changes observed across conditions 280 ms after stimulus onset were reliable. Accordingly, these dominant maps were fitted back to the individual ERP data (by using a spatial fitting procedure, as described above) to estimate their representation across time and conditions. For this purpose, three different time intervals were defined based on the outcome of the spatiotemporal cluster analysis, each lasting 80 ms: a first interval around the peak (as defined using the GFP) of the visual N1 map, a second one corresponding to the occipital P2 map, and a third one immediately following the P2 map, where topographic differences were found by the preceding cluster analysis. For the visual N1 map (136-216 ms post-stimulus onset), a univariate ANOVA performed on the GEV values failed to reveal any significant difference between the four image levels [$F(3, 72) = 0.78$, $p = .507$, $\eta_p^2 = .032$], lending support to the assumption that this first occipital map might correspond to the

early visual encoding or discrimination of the incoming (blurred) stimulus. For the P2 map (216-296 ms post-stimulus onset), the univariate ANOVA showed a significant main effect of image level [$F(3, 72) = 5.04, p = .003, \eta_p^2 = .173$]. Post-hoc comparisons revealed a selective decrease of the global explained variance of this P2 map for the image level corresponding to actual recognition, relative to the three other levels (One image before, $p = .044$; Two images before, $p = .010$; Three image before, $p = .008$). The other pairwise comparisons remained non-significant (all $p_s > .05$). This latter result suggested a substantial decrease of the P2 at the time of recognition, which could reflect either the processing of low-level visual properties of the stimulus (Luck & Hillyard, 1994) or short-term memory load (Wolach & Pratt, 2001), including priming (Gruber & Muller, 2005; Rugg, Soardi, & Doyle, 1995; Wiggs & Martin, 1998). To ascertain the presence of a reliable topographic change during the third time interval (280-360 ms post-stimulus onset), we then submitted the GEV values obtained after fitting during this time interval to a 2 (map configuration: occipito-parietal positive activity vs. occipital negative/frontal positive activity) x 4 (image level: Recognition, One image before, Two images before, Three images before) repeated measures ANOVA. This analysis revealed a main effect of map configuration [$F(1, 18) = 4.89, p = .040, \eta_p^2 = .214$], and a highly significant map configuration x image level interaction [$F(3, 54) = 14.34, p < .001, \eta_p^2 = .443$], corroborating the assumption of a reliable topographic change across the four image levels during this specific time interval, and in sharp contrast to the results obtained for the preceding visual N1 and P2 scalp maps. Interestingly, post-hoc comparisons showed that the GEV of the occipital negative/frontal positive activity progressively increased from three images before recognition to actual recognition (see Figure 3A). More specifically, the GEV of this component was found to reliably increase when comparing actual recognition to one image before [$t(18) = 3.12, p = .006$], two

images before [$t(18) = 2.61, p = .018$] and three images before recognition [$t(18) = 5.02, p < .001$]. Similarly, the substantial increase in GEV as a function of the progressive unfolding of the scene content was evident when comparing one image to three images before recognition [$t(18) = 2.56, p = .020$], as well as two images to three images before recognition [$t(18) = 2.98, p = .008$]. By comparison, the evolution of the GEV values of the concurrent occipito-parietal positive activity over the four image levels showed a different statistical outcome (Figure 3B). T-tests performed on the GEV values extracted for this topographic activity during the same time interval (280-360 ms post-stimulus onset) revealed a significant decrease of the GEV for the image level corresponding to actual recognition, relative to one image [$t(18) = -4.59, p < .001$], two images [$t(18) = -3.65, p = .002$] and three images before recognition [$t(18) = -4.50, p < .001$]. The other pairwise comparisons were not significant. In other words, this significant map configuration x image level interaction indicated a reliable change of microstates during this time interval, when progressively moving from three images before recognition to actual recognition. This topographic change was best explained by a progressive increase of the negative occipital/frontal positive activity when moving closer to recognition, whereas the concurrent occipito-parietal positive activity was found to be stable for the three first levels but showed a sharp decrease at the time corresponding to actual recognition.

We failed to find any significant modulation of these map configurations and their expressions as a function of the emotional content of the scene. We carried out a 2 (map configuration) x 3 (emotional content) x 4 (image level) repeated measures ANOVA on the GEV values extracted during this time interval (280-360 ms post-stimulus onset) that revealed a significant main effect of map configuration [$F(1, 18) = 4.89, p = .040, \eta_p^2 = .213$] and a significant map configuration x image level interaction [$F(3, 54) = 14.35, p < .001, \eta_p^2 = .444$],

but no significant modulation by the emotional content of the scene. This result suggested that these differential neural processes related to visual object recognition were not influenced by the emotional content of the scenes.

Source localization results

Next, we used sLORETA (Pascual-Marqui, 2002) to gain insight into the likely neural generators accounting for the different topographies identified by the previous topographic analyses. More specifically, sLORETA was used to explore the brain regions underlying the substantial topographic change found in the previous analysis, which concerned the 280-360 ms post-stimulus onset interval. Since the change in the electric field configuration was most obvious when directly comparing three images before recognition to actual recognition, we first used this contrast in the inverse solution space and performed statistical non-parametric mapping (SnPM) analyses. For this purpose, amplitude data were first normalized (i.e. total average power equal to unity) prior to performing statistical analyses. This procedure revealed a stronger activation for recognition compared to three images before recognition in the left dACC (-15 \underline{x} , -10 \underline{y} , +45 \underline{z}) [$t(18) = 4.16$, $p < .001$] and right dACC (+15 \underline{x} , -10 \underline{y} , +45 \underline{z}) [$t(18) = 4.50$, $p < .001$], extending bilaterally in the supplementary motor area (SMA; left: -25 \underline{x} , -10 \underline{y} , +45 \underline{z} ; right: +25 \underline{x} , -10 \underline{y} , +45 \underline{z} ; $t(18) = 4.31$, $p < .001$ and $t(18) = 3.01$, $p = .007$, respectively). The reverse contrast (Three images before > Recognition) showed larger activations in the bilateral posterior cingulate cortex (PCC), extending ventrally in the parahippocampal gyrus (PHG) [left: -15 \underline{x} , -65 \underline{y} , +10 \underline{z} ; right: +15 \underline{x} , -65 \underline{y} , +10 \underline{z} ; $t(18) = -4.56$, $p < .001$ and $t(18) = -2.80$, $p = .012$, respectively] (Figure 4A). Next, for each of these regions of interest (ROIs) and each subject separately, we extracted the amplitude value (mean amplitude during the 280-360 ms post-stimulus onset interval) to establish how the activity in these regions actually evolved as a

function of time of recognition. A 2 (side: left vs. right hemisphere) x 2 (ROI: dACC vs. PCC/PHG) x 4 (image level: Recognition, One image before, Two images before, Three images before) repeated measures ANOVA disclosed a significant main effect of ROI [$F(1, 18) = 22.61$, $p < .001$, $\eta_p^2 = .557$] and a marginally significant effect of image level [$F(3, 54) = 2.71$, $p = .074$, $\eta_p^2 = .131$] but, more importantly, a significant ROI x image level interaction [$F(3, 54) = 14.79$, $p < .001$, $\eta_p^2 = .451$]. For the left dACC (Figure 4B), post-hoc comparisons revealed a significantly higher activity during actual recognition, relative to one image [$t(18) = 5.29$, $p < .001$], two images [$t(18) = 5.22$, $p < .001$] or three images before recognition [$t(18) = 4.90$, $p < .007$]. None of the other pairwise comparisons were significant (all $ps > .05$). Inverse solution results obtained for the right dACC showed a very similar outcome: a stronger activity was evidenced when comparing actual recognition to one image [$t(18) = 5.53$, $p < .001$], two images [$t(18) = 4.09$, $p = .001$] or three images before recognition [$t(18) = 4.21$, $p = .001$]. By contrast, statistical analyses performed on the amplitude values extracted from the PCC/PHG showed a different result, mainly characterized by a linear decrease in activity when moving from three images before recognition to actual recognition. For the left PCC/PHG, paired t-tests showed a significantly lower activity in this region during actual recognition, relative to two images [$t(18) = -4.50$, $p < .001$] and three images before recognition [$t(18) = -5.04$, $p < .001$]. Significantly lower neural activity was also evidenced when comparing one image to two images before recognition [$t(18) = -2.12$, $p = .048$], one image to three images before recognition [$t(18) = -3.39$, $p = .003$], and two images to three images before recognition [$t(18) = -2.51$, $p = .022$], suggesting a linear monotonic decrease of activity in this region as a function of accumulation of perceptual evidence (see Figure 4C). The activity extracted in the right PCC/PHG showed a similar trend, although attenuated. Statistical analyses revealed significantly higher amplitude during actual

recognition compared to three images before recognition [$t(18) = -2.31, p = .033$]. The same effect was evidenced when comparing one image to three images before recognition [$t(18) = -2.19, p = .042$].

Finally, we verified whether the activity in these ROIs varied with the emotional content of the scenes or not. In none of the four ROIs did the ANOVA reveal any significant effect of the emotional content of the scene, suggesting that neural processing in these four regions was not influenced by the emotional content of the scene during this specific time interval following stimulus onset, consistent with the topographic analyses reported here above.

Discussion

In this study, we used high density EEG to shed light on the neural events preceding and leading to perceptual decision making during a simple/binary visual categorization task (animacy judgment). For each trial, participants were presented with series of filtered images that were progressively unfolding the content of a complex visual scene, and the participant was asked to discriminate whether this scene contained a living object or not. The sequence started with the presentation of a blurred image, whose content was increasingly revealed by adding up, in a non-linear fashion, HSF information, hence providing a temporal decomposition of a “coarse-to-fine” analysis of the incoming visual stimulus. Since previous models have emphasized such a “coarse-to-fine” analysis subtended by different contributions of low vs. high spatial frequency information (Bar, 2004; Bullier, 2001; Hegdé, 2008), this sequential procedure was then developed to investigate upstream evidence accumulation processes leading to recognition. In addition, the visual scenes were neutral, pleasant or unpleasant, to assess whether their emotional content might affect the expression of neural events foreshadowing overt visual object recognition.

Behavioral results confirmed that this new progressive stimulus revelation task was suited to study the temporal dynamic preceding visual object recognition. Participants consistently waited for sufficient perceptual evidence before categorizing the incoming visual stimulus as either living or non-living with high accuracy. Accurate perceptual decisions (mean % response correct > 87) mainly occurred after the presentation of three (spatial filtering: 64-512 pixels/cycle), four (32-512 pixels/cycle) or five (16-512 pixels/cycle) images, suggesting systematic accumulation of evidence before recognition (see Figure 1B). Consistent with previous studies (Bar, 2004; Delplanque et al., 2007; Schyns & Oliva, 1994), these three image levels presumably contained adequate diagnostic spatial frequency content to perform the animacy judgment task with high accuracy and confidence. Behavioral results obtained for catch trials corroborated the assumption that participants did not simply guess about the content of the scene but, instead, they reliably accumulated perceptual evidence before making a decision about its content, as shown by progressively higher perceptual decisions made during the presentation of the third, fourth and fifth image level. Altogether, these behavioral findings are compatible with predictions arising from accumulator models of decision making, which directly emphasize the accumulation of perceptual evidence in favor of each alternative before a response boundary is surpassed (Gold & Shadlen, 2007; Ratcliff & McKoon, 2008).

Interestingly, our behavioral results also showed reliable differences between the three emotion conditions, indicated by earlier recognitions for unpleasant compared to pleasant scenes. Although speculative, this effect obtained for negatively-valenced scenes might be consistent with a negativity bias effect, as described in the Evaluative Space Model (ESM: Cacioppo, Gardner, & Berntson, 1997; Cacioppo, Gardner, & Berntson, 1999; Cacioppo, & Gardner, 1999; Norris, Gollan, Berntson, & Cacioppo, 2010). This dominant model makes the assumption of

two separable and partially distinct components of the system underlying the evaluation of emotion and affect: (1) positivity, sensitive to appetitive stimuli and promoting approaching behavior; (2) negativity, oriented towards threat or danger, and fostering avoidance (Cacioppo, & Gardner, 1999). These components are characterized by distinctive activation functions, a negativity bias (strongly aversive stimuli elicit stronger responses than appetitive ones) and a positivity offset (when input to the affect system is minimal, positivity outweighs negativity). This negativity bias would lead to slower or diminished responses to non-negative (appetitive/pleasant), relative to negative stimuli (Norris, Gollan, Berntson, & Cacioppo, 2010). However, because neutral scenes were recognized on average earlier than negative scenes, a general negativity bias only does not seem sufficient to account for our behavioral results. Alternatively, we cannot rule out the possibility that picture complexity (or stimulus ambiguity) might actually be different between pleasant, unpleasant and neutral scenes, a factor that could potentially account for the differences in speed of recognition found across these three emotion conditions in our task (see Fig. 1B). However, we first aimed at selecting mildly emotional pictures from the IAPS, which were balanced regarding the animacy dimension but for which clear contrast effects (regarding the valence and arousal dimensions) could be obtained when comparing these neutral, pleasant and unpleasant scenes (see Table 1). This selection procedure resulted in a limited number of individual scenes for each emotion condition. In this context, we could not control for possible low-level visual differences across the three emotion conditions, which may eventually influence performance during the animacy judgment task, even though it remains unclear if a specific low-level visual property may systematically bias the animacy judgments in one direction or the other. Moreover, because the early ERP components (0-200 ms post-stimulus onset, including the N1 and P2) were found to be identical across the three emotion

conditions, a differential recognition effect across the three emotion categories triggered by specific low-level visual properties appears unlikely.

At the electrophysiological level, we found evidence for the involvement of different brain regions that differently contributed to mechanisms of perceptual decision making. More specifically, we looked at electrophysiological effects occurring up to three images before recognition, and eventually found reliable topographic modulations of the ERP signal in a window spanning from 280 to 360 ms after stimulus onset across these four image levels. Notably, two different topographic activities showing different evolution over time (see Figure 3A and 3B) were evidenced. Whereas the explained variance of an occipital negative/frontal positive activity linearly increased from three images before recognition until time of recognition (Figure 3A), the explained variance of a concurrent broad occipito-parietal positive activity remained stable for the different image levels before recognition, but substantially dropped at the time of recognition (Figure 3B). On the other hand, we did not find any significant modulation of these effects by the emotional content of the scene, suggesting a common visual object recognition mechanism for neutral, pleasant and unpleasant pictures during this specific time interval. A possible reason may be that our pre-selected IAPS stimuli were only mildly arousing, in contrast with many previous studies in literature showing reliable visual ERP effects with highly arousing pictures (including the EPN and LPP components, recorded following stimulus onset), relative to neutral, or low arousing scenes (Delplanque et al., 2004; Junghöfer, Bradley, Elbert, & Lang, 2001; Peyk, Schupp, Keil, Elbert, & Junghöfer, 2009; Schupp, Junghöfer, Weike, & Hamm, 2003a; Schupp et al., 2003b; Schupp et al., 2006; Wiens, Peira, Golkar, & Ohman, 2008). In addition, given the specific data analysis used in our study, we cannot rule out

the possibility that these neural processes might be triggered with some delay for emotional compared to neutral pictures, an issue that requires further research.

Since these topographic changes observed at the scalp level necessarily denote alterations in the configuration of the intracranial generators (see Lehmann & Skrandies, 1980; Michel, et al., 1999; Murray, et al., 2008; Pourtois, et al., 2008), we performed complementary source localization analyses to gain insight into the putative configuration of the intracranial generators underlying these two specific topographic activities related to visual object recognition. Inverse solutions based on sLORETA (Pascual-Marqui, 2002) confirmed a substantial shift in the localization of the neural generators underlying these two topographic maps. Comparing three images before recognition to time of recognition revealed a highly significant effect ($p < .001$ corrected) in the PCC/PHG complex (Brodmann's areas, BA 29/30). The reverse contrast revealed a highly significant ($p < .001$ corrected) bilateral effect in the dACC (BA 24) (Figure 4A), with additional effects localized more dorsally, including the SMA bilaterally. This latter effect in the SMA might reflect a motor preparation component that preceded or was associated with the overt recognition of the scene. While the former regions showed a quasi-linear decrease in activation when moving from three images before recognition to actual recognition (see Figure 4C), the extracted activity in the latter region showed instead a stable pattern up to the image level corresponding to actual recognition, where an abrupt and substantial increase was evidenced (see Figure 4B). Hence, distinct effects were found in these two distant regions as a function of accumulation evidence processes preceding actual recognition.

Previous neuroscience studies have linked the PCC, and more specifically its retrosplenial portion (BA 29), to the acquisition of visual memory traces (Dickerson & Eichenbaum, 2010; Shallice et al., 1994; Valenstein et al., 1987). Moreover, recent studies have advocated this brain

region in the formation of canonical representations for typical contexts encountered during visual object recognition (Aminoff, Gronau, & Bar, 2007; Bar, 2009; Bar, Aminoff, Mason, & Fenske, 2007), whereas the adjacent PHG would primarily be involved in the generation of associations related to these contexts (Aminoff et al., 2007; Bar, 2004, 2007; Bar et al., 2006). Therefore, these different brain areas might work in concert and form a functional network involved in the analysis and generation of contextual information presented together with the actual visual object (Bar, 2009). Interestingly, here we found a similar network (involving the PCC and PHG) whose activity was found to monotonically decrease when increasing the spatial frequency content of the input image, and hence when progressively revealing important contextual cues about the object's identity (see Figure 4C). This effect may therefore reflect the processing of diagnostic contextual information (primarily based on LSF information; see Bar, 2009) needed to optimally categorize the content of the image, and eventually perform the animacy judgment task.

By contrast, source localization results showed the involvement of the bilateral dACC (BA 24) in our task, although the response profile in this region (see Figure 4B) was reliably different compared to the PCC/PHG (Figure 4C). The dACC was primarily found to be active during the image level corresponding to the actual recognition of the scene. Activity of the dACC was low and stable during the three image levels before recognition, but substantially increased at the time of recognition, consistent with its involvement in higher-order decision making processes. Various functions have been ascribed to the dACC (Bush, Luu, & Posner, 2000; Ridderinkhof et al., 2004), including reward evaluation (Sanfey, Rilling, Aronson, Nystrom, & Cohen, 2003), response conflict/competition (Botvinick, 2007; Botvinick, Braver, Barch, Carter, & Cohen, 2001; Bush et al., 1998), value judgments (Seitz, Franz, & Azari, 2009), error detection

(Dehaene, Posner, & Tucker, 1994; Holroyd & Coles, 2002) and reward-based decision making (Bush et al., 2002; Hampton & O'Doherty, 2007; Rushworth & Behrens, 2008; Seo & Lee, 2007, 2009; Watanabe, 2007). A critical function of this region may be the monitoring of choice outcomes, with the aim to further adjust sensory acuity in order to improve the organism's response to the environment (Kable & Glimcher, 2009). Recently, the dACC has also been thought to be responsible for both preparation and online adjustments in response to conflicts, considered either as environmental feedbacks or internally generated signals (Anderson, Anderson, Ferris, Fincham, & Jung, 2009; Fincham & Anderson, 2006; Sohn, Albert, Jung, Carter, & Anderson, 2007). Given the rather categorical/all-or-nothing response profile found for the dACC in our study, this effect may therefore reflect the involvement of this region in monitoring the outcome of a decision, which is based in the present case on the rapid and likely incomplete accumulation of perceptual evidence. Importantly, here perceptual evidence remains partial or incomplete at the time of overt decision (Response 1, see Methods). This situation could potentially generate a mild conflict, expressed by the urge to make a rapid decision while the actual accumulation of evidence is not complete yet (Hayden, Pearson, & Platt, 2009; Sarinopoulos et al., 2010; Scheffers & Coles, 2000). Thus, whereas the PCC/PHG complex may be involved in updating contextual information about the scene (based on the rapid extraction and accumulation of LSF information; see Bar, 2009), the dACC may contribute to visual object recognition mechanisms in this task by monitoring the ongoing decision's outcome, and serving therefore as an important interface between the accumulation of perceptual evidence on the one hand, and the implementation and execution of a motor plan on the other hand. The rapid acquisition and accumulation of perceptual/contextual evidence taking place in the PCC/PHG would lead in turn to the monitoring of the actual decision's outcome within the dACC, as well

as SMA to a smaller degree. This rapid transition might occur by the means of reciprocal anatomical connections between these non-overlapping brain regions (Vogt, Finch, & Olson, 1992).

Thus, the experimental design and data analyses used here enabled us to track, using a millisecond time scale, fast evolving upstream perceptual brain processes eventually leading to visual object recognition, whose precise temporal dynamics may be difficult to disclose using the BOLD-fMRI technique only, given its sluggish temporal resolution. However, our ERP are also complementary to previous fMRI results (see Ploran et al., 2007), as they confirm that non-overlapping brain regions (i.e. dorsal ACC vs. PCC/PHG) were reliably active prior to visual object recognition at different latencies (relative to recognition) and with different response profiles, but they also add new critical information about the precise electrophysiological time-course of these upstream perceptual effects, which primarily concerned an interval spanning from 280 to 360 ms post-stimulus onset.

There are a few restrictions to the present study. First, each visual scene was repeated once after a variable and unpredictable time lag, which may have introduced a slight bias in top-down recognition brain mechanisms for repetitions, relative to first presentations of the scenes. However, our ERP analyses failed to reveal any substantial topographical and source localization difference between first presentations and repetitions, suggesting similar accumulation of evidence processes in these two conditions (though occurring earlier for repetitions relative to first presentations). Although this behavioral advantage in speed of recognition for repetitions relative to first presentations does not yield ERP topographic differences in our study, one may assume that additional brain processes may underlie this differential effect, even if they were not visible in the present case. The use of a more explicit (as opposed to implicit) visual encoding

strategy combined with the activation of visual recognition processes based on cues stored in memory might help to reveal different top-down recognition processes for first encounters of the scenes, compared to repetitions. Another limitation concerns the stimuli selected from the IAPS and used in this ERP study. Low-level but uncontrolled differences (e.g. picture complexity, stimulus ambiguity) may exist between neutral, pleasant and unpleasant scenes, which could possibly influence behavioral performance during the animacy judgment task. We cannot exclude the possibility that some of the observed behavioral results (see Fig. 1B) were explained by some uncontrolled “low-level” differences between the three emotion categories (neutral, pleasant and unpleasant scenes).

In sum, our new ERP results provide the first direct electrophysiological evidence of upstream neural events leading to visual object recognition, highlighting distinct effects in the PCC/PHG and dACC during this process rapidly following stimulus onset. Our findings are also in line with recent fMRI results showing that these two regions may differentially contribute to mechanisms of perceptual decision making (Ploran et al., 2007; Wheeler et al., 2008). In addition, the use of high density scalp EEG helped us shed light on the distinctive temporal contributions of these two regions during visual object recognition. Whereas the PCC/PHG complex was involved in the accumulation of perceptual evidence prior to actual recognition, the dACC was likely implicated at a later stage in the monitoring of the decision’s outcome (see also Ploran, et al., 2007). Future studies are needed to establish whether these two distant regions may exhibit some coupling or reciprocal interactions during upstream perceptual processes leading to visual object recognition.

Acknowledgements

This work is supported by grants from the European Research Council (Starting Grant #200758) and Ghent University (BOF Grant #05Z01708).

References

- Agresti, A. (2007). *An Introduction to Categorical Data Analysis*. New York: Wiley.
- Aminoff, E., Gronau, N., & Bar, M. (2007). The parahippocampal cortex mediates spatial and nonspatial associations. *Cerebral Cortex*, *17*(7), 1493-1503.
- Anderson, J. R., Anderson, J. F., Ferris, J. L., Fincham, J. M., & Jung, K. J. (2009). Lateral inferior prefrontal cortex and anterior cingulate cortex are engaged at different stages in the solution of insight problems. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(26), 10799-10804.
- Bar, M. (2003). A cortical mechanism for triggering top-down facilitation in visual object recognition. *Journal of Cognitive Neuroscience*, *15*(4), 600-609.
- Bar, M. (2004). Visual objects in context. *Nature Reviews Neuroscience*, *5*(8), 617-629.
- Bar, M. (2007). The proactive brain: using analogies and associations to generate predictions. *Trends Cogn Sci*, *11*(7), 280-289.
- Bar, M. (2009). The proactive brain: memory for predictions. *Philosophical Transactions of the Royal Society B-Biological Sciences*, *364*(1521), 1235-1243.
- Bar, M., Aminoff, E., Mason, M., & Fenske, M. (2007). The units of thought. *Hippocampus*, *17*(6), 420-428.
- Bar, M., Kassam, K. S., Ghuman, A. S., Boshyan, J., Schmidt, A. M., Dale, A. M., et al. (2006). Top-down facilitation of visual recognition. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(2), 449-454.
- Biederman, I. (1987). Recognition-by-Components: A Theory of Human Image Understanding. *Psychological Review*, *94*(2), 115-147.
- Biederman, I., & Bar, M. (1999). One-shot viewpoint invariance in matching novel objects. *Vision Research*, *39*(17), 2885-2899.
- Blair, M. R., Watson, M. R., Walshe, R. C., & Maj, F. (2009). Extremely Selective Attention: Eye-Tracking Studies of the Dynamic Allocation of Attention to Stimulus Features in Categorization. *Journal of Experimental Psychology-Learning Memory and Cognition*, *35*(5), 1196-1206.
- Botvinick, M. M. (2007). Conflict monitoring and decision making: Reconciling two perspectives on anterior cingulate function. *Cognitive Affective & Behavioral Neuroscience*, *7*(4), 356-366.
- Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychological Review*, *108*(3), 624-652.
- Bradley, M. M., & Lang, P. J. (1994). Measuring Emotion: The Self-Assessment Mannequin and the Semantic Differential. *Journal of Behavior Therapy and Experimental Psychiatry*, *25*(1), 49-59.
- Bullier, J. (2001). Integrated model of visual processing. *Brain Research Reviews*, *36*(2-3), 96-107.
- Bush, G., Luu, P., & Posner, M. I. (2000). Cognitive and emotional influences in anterior cingulate cortex. *Trends in Cognitive Sciences*, *4*(6), 215-222.
- Bush, G., Vogt, B. A., Holmes, J., Dale, A. M., Greve, D., Jenike, M. A., et al. (2002). Dorsal anterior cingulate cortex: A role in reward-based decision making. *Proceedings of the National Academy of Sciences of the United States of America*, *99*(1), 523-528.

- Bush, G., Whalen, P. J., Rosen, B. R., Jenike, M. A., McInerney, S. C., & Rauch, S. L. (1998). The counting stroop: An interference task specialized for functional neuroimaging - Validation study with functional MRI. *Human Brain Mapping, 6*(4), 270-282.
- Cacioppo, J. T., & Gardner, W. L. (1999). Emotion. *Annual Review of Psychology, 50*, 191-214.
- Cacioppo, J. T., Gardner, W. L., & Berntson, G. G. (1997). Beyond Bipolar Conceptualizations and Measures: The Case of Attitudes and Evaluative Space. *Personality and Social Psychology Review, 1*(1), 3-25.
- Cacioppo, J. T., Gardner, W. L., & Berntson, G. G. (1999). The affect system has parallel and integrative processing components: Form follows function. *Journal of Personality and Social Psychology, 76*(5), 839-855.
- Carlson, T., Grol, M. J., & Verstraten, F. A. J. (2006). Dynamics of visual recognition revealed by fMRI. *Neuroimage, 32*(2), 892-905.
- Carretié, L., Hinojosa, J. A., Martin-Loeches, M., Mercado, F., & Tapia, M. (2004). Automatic attention to emotional stimuli: Neural correlates. *Human Brain Mapping, 22*(4), 290-299.
- Crowley, K. E., & Colrain, I. M. (2004). A review of the evidence for P2 being an independent component process: age, sleep and modality. *Clinical Neurophysiology, 115*(4), 732-744.
- D'Hondt, F., Lassonde, M., Collignon, O., Dubarry, A. S., Robert, M., Rigoulot, S., et al. Early brain-body impact of emotional arousal. *Frontiers in Human Neuroscience, 4*, 33.
- Dehaene, S., Posner, M. I., & Tucker, D. M. (1994). Localization of a Neural System for Error-Detection and Compensation. *Psychological Science, 5*(5), 303-305.
- Delplanque, S., Lavoie, M. E., Hot, P., Silvert, L., & Sequeira, H. (2004). Modulation of cognitive processing by emotional valence studied through event-related potentials in humans. *Neuroscience Letters, 356*(1), 1-4.
- Delplanque, S., N'diaye, K., Scherer, K., & Grandjean, D. (2007). Spatial frequencies or emotional effects? A systematic measure of spatial frequencies for IAPS pictures by a discrete wavelet analysis. *Journal of Neuroscience Methods, 165*(1), 144-150.
- Di Lollo, V., Enns, J. T., & Rensink, R. A. (2000). Competition for consciousness among visual events: The psychophysics of reentrant visual processes. *Journal of Experimental Psychology-General, 129*(4), 481-507.
- Dickerson, B. C., & Eichenbaum, H. (2010). The Episodic Memory System: Neurocircuitry and Disorders. *Neuropsychopharmacology, 35*(1), 86-104.
- Egner, T., & Hirsch, J. (2005). Cognitive control mechanisms resolve conflict through cortical amplification of task-relevant information. *Nature Neuroscience, 8*(12), 1784-1790.
- Enns, J. T., & Lleras, A. (2008). What's next? New evidence for prediction in human vision. *Trends in Cognitive Sciences, 12*(9), 327-333.
- Fincham, J. M., & Anderson, J. R. (2006). Distinct roles of the anterior cingulate and prefrontal cortex in the acquisition and performance of a cognitive skill. *Proceedings of the National Academy of Sciences of the United States of America, 103*(34), 12941-12946.
- Freunberger, R., Klimesch, W., Doppelmayr, M., & Holler, Y. (2007). Visual P2 component is related to theta phase-locking. *Neuroscience Letters, 426*(3), 181-186.
- Friston, K. (2005). A theory of cortical responses. *Philosophical Transactions of the Royal Society B-Biological Sciences, 360*(1456), 815-836.
- Friston, K., & Kiebel, S. (2009). Predictive coding under the free-energy principle. *Philosophical Transactions of the Royal Society B-Biological Sciences, 364*(1521), 1211-1221.

- Fuchs, M., Kastner, J., Wagner, M., Hawes, S., & Ebersole, J. S. (2002). A standardized boundary element method volume conductor model. *Clinical Neurophysiology*, *113*(5), 702-712.
- Ganis, G., Schendan, H. E., & Kosslyn, S. M. (2007). Neuroimaging evidence for object model verification theory: Role of prefrontal control in visual object categorization. *Neuroimage*, *34*(1), 384-398.
- Gold, J. I., & Shadlen, M. N. (2007). The neural basis of decision making. *Annual Review of Neuroscience*, *30*, 535-574.
- Gratton, G., Coles, M. G. H., & Donchin, E. (1983). A New Method for Off-Line Removal of Ocular Artifact. *Electroencephalography and Clinical Neurophysiology*, *55*(4), 468-484.
- Grossberg, S. (2009). Cortical and subcortical predictive dynamics and learning during perception, cognition, emotion and action. *Philosophical Transactions of the Royal Society B-Biological Sciences*, *364*(1521), 1223-1234.
- Gruber, T., & Muller, M. M. (2005). Oscillatory brain activity dissociates between associative stimulus content in a repetition priming task in the human EEG. *Cerebral Cortex*, *15*(1), 109-116.
- Hampton, A. N., & O'Doherty, J. P. (2007). Decoding the neural substrates of reward-related decision making with functional MRI. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(4), 1377-1382.
- Hanes, D. P., & Schall, J. D. (1996). Neural control of voluntary movement initiation. *Science*, *274*(5286), 427-430.
- Hayden, B. Y., Pearson, J. M., & Platt, M. L. (2009). Fictive Reward Signals in the Anterior Cingulate Cortex. *Science*, *324*(5929), 948-950.
- Heekeren, H. R., Marrett, S., Bandettini, P. A., & Ungerleider, L. G. (2004). A general mechanism for perceptual decision-making in the human brain. *Nature*, *431*(7010), 859-862.
- Heekeren, H. R., Marrett, S., & Ungerleider, L. G. (2008). The neural systems that mediate human perceptual decision making. *Nature Reviews Neuroscience*, *9*(6), 467-479.
- Hegd , J. (2008). Time course of visual perception: Coarse-to-fine processing and beyond. *Progress in Neurobiology*, *84*(4), 405-439.
- Holroyd, C. B., & Coles, M. G. H. (2002). The neural basis of human error processing: Reinforcement learning, dopamine, and the error-related negativity. *Psychological Review*, *109*(4), 679-709.
- James, T. W., Humphrey, G. K., Gati, J. S., Menon, R. S., & Goodale, M. A. (2000). The effects of visual object priming on brain activation before and after recognition. *Current Biology*, *10*(17), 1017-1024.
- Jungh fer, M., Bradley, M. M., Elbert, T. R., & Lang, P. J. (2001). Fleeting images: A new look at early emotion discrimination. *Psychophysiology*, *38*(2), 175-178.
- Jurcak, V., Tsuzuki, D., & Dan, I. (2007). 10/20, 10/10, and 10/5 systems revisited: Their validity as relative head-surface-based positioning systems. *Neuroimage*, *34*(4), 1600-1611.
- Kable, J. W., & Glimcher, P. W. (2009). The Neurobiology of Decision: Consensus and Controversy. *Neuron*, *63*(6), 733-745.
- Kersten, D., Mamassian, P., & Yuille, A. (2004). Object perception as Bayesian inference. *Annual Review of Psychology*, *55*, 271-304.

- Kim, J. N., & Shadlen, M. N. (1999). Neural correlates of a decision in the dorsolateral prefrontal cortex of the macaque. *Nature Neuroscience*, 2(2), 176-185.
- Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (2005). *International Affective Picture System (IAPS): Digitized photographs, instruction manual, and affective ratings (Tech. Rep. No. A-6)*. Gainesville: University of Florida, Center for Research in Psychophysiology.
- Lehmann, D., & Skrandies, W. (1980). Reference-Free Identification of Components of Checkerboard-Evoked Multichannel Potential Fields. *Electroencephalography and Clinical Neurophysiology*, 48(6), 609-621.
- Lithari, C., Frantzidis, C. A., Papadelis, C., Vivas, A. B., Klados, M. A., Kourtidou-Papadeli, C., et al. (2010). Are Females More Responsive to Emotional Stimuli? A Neurophysiological Study Across Arousal and Valence Dimensions. *Brain Topography*, 23(1), 27-40.
- Lomber, S. G. (2002). Learning to see the trees before the forest: Reversible deactivation of the superior colliculus during learning of local and global visual features. *Proceedings of the National Academy of Sciences of the United States of America*, 99(6), 4049-4054.
- Luck, S. J., & Hillyard, S. A. (1994). Electrophysiological Correlates of Feature Analysis during Visual-Search. *Psychophysiology*, 31(3), 291-308.
- Maio, G. R., & Esses, V. M. (2001). The need for affect: Individual differences in the motivation to approach or avoid emotions. *Journal of Personality*, 69(4), 583-615.
- Marr, D. (1982). *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information*. New York: Freeman.
- Michel, C. M., Seeck, M., & Landis, T. (1999). Spatiotemporal dynamics of human cognition. *News in Physiological Sciences*, 14, 206-214.
- Michel, C. M., Thut, G., Morand, S., Khateb, A., Pegna, A. J., de Peralta, R. G., et al. (2001). Electric source imaging of human brain functions. *Brain Research Reviews*, 36(2-3), 108-118.
- Murray, M. M., Brunet, D., & Michel, C. M. (2008). Topographic ERP analyses: A step-by-step tutorial review. *Brain Topography*, 20(4), 249-264.
- Navon, D. (1977). Forest Before Trees: The Precedence of Global Features in Visual Perception. *Cognitive Psychology*, 9(3), 353-383.
- Norris, C. J., Gollan, J., Berntson, G. G., & Cacioppo, J. T. (2010). The current status of research on the structure of evaluative space. *Biological Psychology*, 84(3), 422-436.
- Oliva, A., & Torralba, A. (2007). The role of context in object recognition. *Trends in Cognitive Sciences*, 11(12), 520-527.
- Olofsson, J. K., Nordin, S., Sequeira, H., & Polich, J. (2008). Affective picture processing: An integrative review of ERP findings. *Biological Psychology*, 77(3), 247-265.
- Opris, I., & Bruce, C. J. (2005). Neural circuitry of judgment and decision mechanisms. *Brain Research Reviews*, 48(3), 509-526.
- Palmer, S. E. (1999). *Vision Science: Photons to Phenomenology*. Cambridge, MA: MIT Press.
- Pascual-Marqui, R. D. (2002). Standardized low-resolution brain electromagnetic tomography (sLORETA): Technical details. *Methods and Findings in Experimental and Clinical Pharmacology*, 24, 5-12.
- Pascual-Marqui, R. D., Michel, C. M., & Lehmann, D. (1995). Segmentation of Brain Electrical Activity into Microstates: Model Estimation and Validation. *Ieee Transactions on Biomedical Engineering*, 42(7), 658-665.

- Perrin, F., Pernier, J., Bertrand, O., & Echallier, J. F. (1989). Spherical Splines for Scalp Potential and Current-Density Mapping. *Electroencephalography and Clinical Neurophysiology*, 72(2), 184-187.
- Peyk, P., Schupp, H. T., Keil, A., Elbert, T., & Junghöfer, M. (2009). Parallel processing of affective visual stimuli. *Psychophysiology*, 46(1), 200-208.
- Peyrin, C., Michel, C. M., Schwartz, S., Thut, G., Seghier, M., Landis, T., et al. (2010). *The neural substrates and timing of top-down processes during coarse-to-fine categorization of visual scenes: a combined fMRI and ERP study* (Vol. 22).
- Picton, T. W., Bentin, S., Berg, P., Donchin, E., Hillyard, S. A., Johnson, R., et al. (2000). Guidelines for using human event-related potentials to study cognition: Recording standards and publication criteria. *Psychophysiology*, 37(2), 127-152.
- Ploran, E. J., Nelson, S. M., Velanova, K., Donaldson, D. I., Petersen, S. E., & Wheeler, M. E. (2007). Evidence accumulation and the moment of recognition: Dissociating perceptual recognition processes using fMRI. *Journal of Neuroscience*, 27(44), 11912-11924.
- Pourtois, G., Dan, E. S., Grandjean, D., Sander, D., & Vuilleumier, P. (2005). Enhanced extrastriate visual response to bandpass spatial frequency filtered fearful faces: Time course and topographic evoked-potentials mapping. *Human Brain Mapping*, 26(1), 65-79.
- Pourtois, G., Delplanque, S., Michel, C., & Vuilleumier, P. (2008). Beyond conventional event-related brain potential (ERP): Exploring the time-course of visual emotion processing using topographic and principal component analyses. *Brain Topography*, 20(4), 265-277.
- Proverbio, A. M., Adorni, R., Zani, A., & Trestianu, L. (2009). Sex differences in the brain response to affective scenes with or without humans. *Neuropsychologia*, 47(12), 2374-2388.
- Ranganath, C., Cohen, M. X., Dam, C., & D'Esposito, M. (2004). Inferior temporal, prefrontal, and hippocampal contributions to visual working memory maintenance and associative memory retrieval. *Journal of Neuroscience*, 24(16), 3917-3925.
- Ranganath, C., & Rainer, G. (2003). Neural mechanisms for detecting and remembering novel events. *Nature Reviews Neuroscience*, 4(3), 193-202.
- Rao, R. P. N., & Ballard, D. H. (1999). Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nature Neuroscience*, 2(1), 79-87.
- Ratcliff, R. (1978). Theory of Memory Retrieval. *Psychological Review*, 85(2), 59-108.
- Ratcliff, R., & McKoon, G. (1982). Speed and accuracy in the processing of false statements about semantic information. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 8, 16-36.
- Ratcliff, R., & McKoon, G. (2008). The diffusion decision model: Theory and data for two-choice decision tasks. *Neural Computation*, 20(4), 873-922.
- Ratcliff, R., Thapar, A., Gomez, P., & McKoon, G. (2004). A diffusion model analysis of the effects of aging in the lexical-decision task. *Psychol Aging*, 19(2), 278-289.
- Ridderinkhof, K. R., Ullsperger, M., Crone, E. A., & Nieuwenhuis, S. (2004). The role of the medial frontal cortex in cognitive control. *Science*, 306(5695), 443-447.
- Riesenhuber, M., & Poggio, T. (1999). Hierarchical models of object recognition in cortex. *Nature Neuroscience*, 2(11), 1019-1025.

- Roets, A., & Van Hiel, A. (2007). Separating ability from need: Clarifying the dimensional structure of the need for closure scale. *Personality and Social Psychology Bulletin*, *33*(2), 266-280.
- Rugg, M. D., Soardi, M., & Doyle, M. C. (1995). Modulation of event-related potentials by the repetition of drawings of novel objects. *Cognitive Brain Research*, *3*(1), 17-24.
- Rushworth, M. F. S., & Behrens, T. E. J. (2008). Choice, uncertainty and value in prefrontal and cingulate cortex. *Nature Neuroscience*, *11*(4), 389-397.
- Sabatinelli, D., Flaisch, T., Bradley, M. M., Fitzsimmons, J. R., & Lang, P. J. (2004). Affective picture perception: gender differences in visual cortex? *Neuroreport*, *15*(7), 1109-1112.
- Sabatinelli, D., Lang, P. J., Bradley, M. M., Costa, V. D., & Keil, A. (2009). The timing of emotional discrimination in human amygdala and ventral visual cortex. *J Neurosci*, *29*(47), 14864-14868.
- Sanfey, A. G., Loewenstein, G., McClure, S. M., & Cohen, J. D. (2006). Neuroeconomics: cross-currents in research on decision-making. *Trends in Cognitive Sciences*, *10*(3), 108-116.
- Sanfey, A. G., Rilling, J. K., Aronson, J. A., Nystrom, L. E., & Cohen, J. D. (2003). The neural basis of economic decision-making in the ultimatum game. *Science*, *300*(5626), 1755-1758.
- Sanocki, T. (1993). Time-Course of Object Identification - Evidence for a Global-to-Local Contingency. *Journal of Experimental Psychology-Human Perception and Performance*, *19*(4), 878-898.
- Sarinopoulos, I., Grupe, D. W., Mackiewicz, K. L., Herrington, J. D., Lor, M., Steege, E. E., et al. (2010). Uncertainty during Anticipation Modulates Neural Responses to Aversion in Human Insula and Amygdala. *Cerebral Cortex*, *20*(4), 929-940.
- Scheffers, M. K., & Coles, M. G. H. (2000). Performance monitoring in a confusing world: Error-related brain activity, judgments of response accuracy, and types of errors. *Journal of Experimental Psychology-Human Perception and Performance*, *26*(1), 141-151.
- Schupp, H. T., Cuthbert, B. N., Bradley, M. M., Cacioppo, J. T., Ito, T., & Lang, P. J. (2000). Affective picture processing: The late positive potential is modulated by motivational relevance. *Psychophysiology*, *37*(2), 257-261.
- Schupp, H. T., Junghöfer, M., Weike, A. I., & Hamm, A. O. (2003a). Attention and emotion: an ERP analysis of facilitated emotional stimulus processing. *Neuroreport*, *14*(8), 1107-1110.
- Schupp, H. T., Junghöfer, M., Weike, A. I., & Hamm, A. O. (2003b). Emotional facilitation of sensory processing in the visual cortex. *Psychological Science*, *14*(1), 7-13.
- Schupp, H. T., Stockburger, J., Codispoti, M., Junghöfer, M., Weike, A. I., & Hamm, A. O. (2006). Stimulus novelty and emotion perception: the near absence of habituation in the visual cortex. *Neuroreport*, *17*(4), 365-369.
- Schyns, P. G., & Oliva, A. (1994). From Blobs to Boundary Edges: Evidence for Time- and Spatial-Scale-Dependent Scene Recognition. *Psychological Science*, *5*(4), 195-200.
- Seitz, R. J., Franz, M., & Azari, N. P. (2009). Value judgments and self-control of action: The role of the medial frontal cortex. *Brain Research Reviews*, *60*(2), 368-378.
- Sekihara, K., Sahani, M., & Nagarajan, S. S. (2005). Localization bias and spatial resolution of adaptive and non-adaptive spatial filters for MEG source reconstruction. *Neuroimage*, *25*(4), 1056-1067.
- Seo, H., & Lee, D. (2007). Temporal filtering of reward signals in the dorsal anterior cingulate cortex during a mixed-strategy game. *Journal of Neuroscience*, *27*(31), 8366-8377.

- Seo, H., & Lee, D. (2009). Behavioral and Neural Changes after Gains and Losses of Conditioned Reinforcers. *Journal of Neuroscience*, 29(11), 3627-3641.
- Serences, J. T. (2008). Value-Based Modulations in Human Visual Cortex. *Neuron*, 60(6), 1169-1181.
- Shadlen, M. N., & Newsome, W. T. (2001). Neural basis of a perceptual decision in the parietal cortex (area LIP) of the rhesus monkey. *Journal of Neurophysiology*, 86(4), 1916-1936.
- Shallice, T., Fletcher, P., Frith, C. D., Grasby, P., Frackowiak, R. S. J., & Dolan, R. J. (1994). Brain Regions Associated with Acquisition and Retrieval of Verbal Episodic Memory. *Nature*, 368(6472), 633-635.
- Smith, P. L., & Ratcliff, R. (2004). Psychology and neurobiology of simple decisions. *Trends in Neurosciences*, 27(3), 161-168.
- Sohn, M. H., Albert, M. V., Jung, K. J., Carter, C. S., & Anderson, J. R. (2007). Anticipation of conflict monitoring in the anterior cingulate cortex and the prefrontal cortex. *Proceedings of the National Academy of Sciences of the United States of America*, 104(25), 10330-10334.
- Summerfield, C., & Egner, T. (2009). Expectation (and attention) in visual cognition. *Trends in Cognitive Sciences*, 13(9), 403-409.
- Tanaka, K. (1996). Inferotemporal cortex and object vision. *Annual Review of Neuroscience*, 19, 109-139.
- Treisman, A. M., & Kanwisher, N. G. (1998). Perceiving visually presented objects: recognition, awareness, and modularity. *Current Opinion in Neurobiology*, 8(2), 218-226.
- Ullsperger, M., Volz, K. G., & von Cramon, D. Y. (2004). A common neural system signaling the need for behavioral changes. *Trends in Cognitive Sciences*, 8(10), 445-446.
- Ungerleider, L. G., & Mishkin, M. (1982). Two cortical visual systems. In D. J. Ingle, M. A. Goodale & R. J. W. Mansfield (Eds.), *Analysis of visual behavior*. Cambridge, MA: MIT Press.
- Valenstein, E., Bowers, D., Verfaellie, M., Heilman, K. M., Day, A., & Watson, R. T. (1987). Retrosplenial Amnesia. *Brain*, 110, 1631-1646.
- Van der Ploeg, H., Defares, P. B., & Spielberger, C. D. (1979). *Zelfbeoordelingsvragenlijst STAI, versie DY-1 en DY-2 [Self report questionnaire STAI, version DY-1 and DY-2]*. Lisse, the Netherlands: Swets & Zeitlinger.
- Vogel, E. K., & Luck, S. J. (2000). The visual N1 component as an index of a discrimination process. *Psychophysiology*, 37(2), 190-203.
- Vogt, B. A., Finch, D. M., & Olson, C. R. (1992). Functional Heterogeneity in Cingulate Cortex: The Anterior Executive and Posterior Evaluative Regions. *Cerebral Cortex*, 2(6), 435-443.
- Vuilleumier, P. (2005). How brains beware: neural mechanisms of emotional attention. *Trends Cogn Sci*, 9(12), 585-594.
- Vuilleumier, P., Armony, J. L., Driver, J., & Dolan, R. J. (2003). Distinct spatial frequency sensitivities for processing faces and emotional expressions. *Nature Neuroscience*, 6(6), 624-631.
- Watanabe, M. (2007). Role of anticipated reward in cognitive behavioral control. *Current Opinion in Neurobiology*, 17(2), 213-219.
- Wheeler, M. E., Petersen, S. E., Nelson, S. M., Ploran, E. J., & Velanova, K. (2008). Dissociating Early and Late Error Signals in Perceptual Recognition. *Journal of Cognitive Neuroscience*, 20(12), 2211-2225.

- Wiens, S., Peira, N., Golkar, A., & Ohman, A. (2008). Recognizing Masked Threat: Fear Betrays, But Disgust You Can Trust. *Emotion*, 8(6), 810-819.
- Wiggs, C. L., & Martin, A. (1998). Properties and mechanisms of perceptual priming. *Current Opinion in Neurobiology*, 8(2), 227-233.
- Wolach, I., & Pratt, H. (2001). The mode of short-term memory encoding as indicated by event-related potentials in a memory scanning task with distractions. *Clinical Neurophysiology*, 112(1), 186-197.
- Yuille, A., & Kersten, D. (2006). Vision as Bayesian inference: analysis by synthesis? *Trends in Cognitive Sciences*, 10(7), 301-308.

Footnote

¹IAPS image codes. WOMEN PICTURE SET: practice (living) 2220, 2635, 4631, 4651, 4669; practice (non-living) 7002, 7009, 7161, 7590, 7820; neutral (living) 1616, 2190, 2191, 2381, 2383, 2385, 2393, 2480, 2495, 2570, 2595, 2749, 2840, 2890, 4250, 4255, 4310, 9070, 3550.2, 6570.2, 2516, 2702, 8192; neutral (non-living) 6150, 7006, 7010, 7020, 7031, 7035, 7036, 7041, 7050, 7130, 7160, 7175, 7179, 7185, 7187, 7217, 7233, 7235, 7830, 7950, 7025, 7500, 7705; negative (living) 1052, 1201, 1525, 1932, 2276, 2490, 2694, 2715, 2753, 3181, 3300, 4621, 6311, 9041, 9046, 9331, 9404, 9417, 2055.1, 2900.1, 1274, 9160, 9592; negative (non-living) 2692, 5971, 6020, 6230, 6241, 6610, 6800, 9000, 9280, 9320, 9340, 9373, 9470, 9471, 9495, 9611, 9620, 9622, 9630, 9830, 9001, 9290, 9621; positive (living) 1604, 1610, 1721, 2209, 2345, 4510, 4538, 4572, 4626, 4640, 4660, 5621, 8034, 8041, 8080, 8200, 8370, 8470, 8490, 8496, 1740, 4531, 8330; positive (non-living) 2791, 5220, 5450, 5480, 5551, 5594, 5600, 5779, 5780, 5891, 5982, 5994, 7200, 7280, 7350, 7470, 7480, 7545, 7580, 8510, 5300, 7390, 8502. MEN PICTURE SET: practice (living) 1112, 2210, 2214, 2393, 3210; practice (non-living) 7030, 7035, 7050, 7185, 7235; neutral (living) 1101, 1230, 1390, 1935, 1945, 2005, 2220, 2441, 2487, 2514, 2516, 2690, 2749, 2830, 2870, 4503, 4520, 4532, 9700, 2745.1, 1310, 1321, 2635; neutral (non-living) 5390, 5731, 6150, 6800, 7009, 7010, 7020, 7036, 7038, 7041, 7090, 7160, 7161, 7179, 7184, 7186, 7207, 7211, 7233, 7283, 7002, 7100, 7285; negative (living) 2053, 2095, 2141, 2710, 2750, 3181, 6243, 6311, 6312, 6315, 6510, 6530, 6821, 6838, 7380, 9250, 9500, 2352.2, 3550.1, 6570.1, 2683, 2900, 9180; negative (non-living) 6260, 6300, 9000, 9008, 9010, 9090, 9280, 9290, 9301, 9320, 9340, 9470, 9471, 9600, 9611, 9620, 9630, 9830, 9911, 9912, 9001, 9360, 9621; positive (living) 1463, 1811, 1999, 2057, 2208, 2209, 2216, 4001, 4599, 4641, 4653, 4656, 4676, 4681, 4689, 4810, 8116, 8120, 8200, 8496, 1460, 2340, 8470; positive (non-

living) 5260, 5270, 5480, 5594, 5600, 5660, 5700, 5760, 5780, 5910, 5982, 7200, 7230, 7270,
7330, 7480, 7580, 8170, 8502, 8510, 7260, 7350, 7470.

Figure caption

Figure 1: (A) Example of stimuli, and progressive unfolding task. Neutral, unpleasant and pleasant IAPS scenes (not shown here for copyright reasons) were presented to participants in random order. For each trial, the content of the scene was progressively revealed in six successive steps by adding high spatial frequency information to the initial blurred (meaningless) picture. Each image level was presented for 500 ms, followed by a 250ms blank screen. Participants were required to press the spacebar key as soon as they could decide whether the scene contained a living object or not, thereby interrupting the sequence (Response1). Next, they validated their response choice by pressing one out of two pre-defined response keys (Response2). (B) Cumulative percentages of correct Response1 as a function of the six image levels, separately for each emotion category. Vertical bars correspond to standard errors of the means. The results showed that participants were able to make the animacy judgment task reliably earlier (i.e. less perceptual evidence needed) for neutral (solid line), relative to unpleasant (dotted line) or pleasant (dashed line) scenes. The shift of the psychometric function for unpleasant compared to pleasant scenes was also significant (see behavioral results).

Figure 2: (A) Grand-average ($N = 19$) ERP waveforms (obtained for the image level corresponding to actual recognition) for the 128 electrodes (butterfly). The red dashed vertical line indicates the onset of the visual stimulus. Clear exogenous N1 and P2 ERP components were recorded after stimulus onset. (B) Results of the spatiotemporal cluster analysis (from stimulus onset until 600 ms after stimulus onset) for the four main conditions (Recognition, One image before, Two images before and Three images before recognition). A main solution with five dominant topographic maps was found to explain $> 90\%$ of the variance. Scalp topographies of

the N1 and P2 were shared across the four conditions, suggesting a similar low-level encoding and processing of the visual stimulus. By contrast, starting 280 ms after stimulus onset, reliable topographic alterations were found between the four image levels. Whereas a broad occipito-parietal activity was generated following the P2 for all the image levels preceding recognition (in red), a distinctive occipital negative component (with a concurrent positive, frontal counterpart) was generated at the same latency for the image level corresponding to actual recognition (in green). (C) Horizontal, frontal and occipital views of the five dominant maps (including the N1 and P2) found in the spatial cluster analysis. Amplitude differences were normalized (i.e. the amplitude value at each electrode was divided by the GFP).

Figure 3: (A) Statistical results (* $p < .05$; ** $p < .01$; *** $p < .001$; vertical bars correspond to standard errors of the means), obtained after the fitting procedure, for the dominant topography characterized by an occipital negativity and a frontal positivity showed a linear increase when moving closer to recognition. The Global Explained Variance (GEV, arbitrary units) was computed during an 80ms time interval (280-360 ms post-stimulus onset) and is presented separately for the four conditions. This analysis showed a linear increase of the GEV when moving from three images before recognition to actual recognition. (B) Results obtained for the concurrent topography of the occipito-parietal positivity identified during the same time interval. Unlike the occipital negative/frontal positive activity, this scalp configuration showed an abrupt decrease for the image level corresponding to recognition, relative to the three preceding images levels where the variance remained stable.

Figure 4: Source localization results. (A) Comparing actual recognition to three images before recognition during the 280-360 ms post-stimulus onset interval disclosed a highly significant effect ($p < .001$ corrected) in the dorsal ACC, bilaterally ($\pm 15x$, $+10y$, $+45z$). The reverse contrast revealed a highly significant effect ($p < .001$ corrected) in the posterior cingulate cortex, extending ventrally towards the parahippocampal gyrus ($\pm 15x$, $-65y$, $+10z$). A: anterior; P: posterior; L: left; R: right. (B) Mean activity extracted from the left dorsal ACC as a function of image level (* $p < .05$; ** $p < .01$; *** $p < .001$). Vertical bars correspond to standard errors of the means. In this dorsal ACC region, a sharp increase was found for recognition, relative to the three preceding levels. (C) By contrast, in the left PPC, a monotonic linear decrease of activity was evidenced when moving towards recognition.

Table 1


Mean values and standard deviations (in parenthesis) of valence and arousal scores of the selected IAPS pictures

Picture set	Valence	Arousal
Men		
<u>Neutral</u>	5.12 (1.29)	3.90 (1.88)
<u>Unpleasant</u>	3.12 (1.62)	4.93 (2.14)
<u>Pleasant</u>	7.31 (1.57)	5.27 (2.30)
Women		
<u>Neutral</u>	4.96 (1.30)	3.34 (1.96)
<u>Unpleasant</u>	2.75 (1.62)	5.27 (2.19)
<u>Pleasant</u>	7.31 (1.60)	4.92 (2.37)

Note. Scores range from 1 to 9. Independent samples t-test confirmed no significant difference between women's and men's picture sets, both for valence [$t(274) = -0.31, p = .760$] and arousal [$t(274) = -1.18, p = .238$] scores. In addition, no difference was found between pictures containing living [valence: $t(136) = 0.10, p = .920$; arousal: $t(136) = -0.71, p = .478$] and non-living [valence: $t(136) = -0.54, p = .590$; arousal: $t(136) = -1.01, p = .315$] objects.

Table 2

The six frequency bands used to filter the IAPS images selected in our study

Image Level	Frequency Band (pixels/cycle)	
1	256-512	LSF
2	128-512	
3	64-512	
4	32-512	
5	16-512	
6	0-512	

Note. HSF: high spatial frequencies; LSF: low spatial frequencies.

Table 3

Mean values and standard deviations (in parenthesis) of the scores obtained for each questionnaire (and relative subscales) administered at the end of the experiment

Questionnaire	Score
STAI-T	35.58 (8.73)
NFCS	3.57 (0.60)
<u>Order</u>	3.83 (0.83)
<u>Predictability</u>	3.51 (1.07)
<u>Decisiveness</u>	3.56 (0.60)
<u>Ambiguity</u>	3.75 (0.80)
<u>Closedmindedness</u>	3.12 (0.50)
NFAS	3.71 (0.55)
<u>Approach</u>	4.27 (0.77)
<u>Avoidance</u>	3.15 (1.28)

Note. STAI-T: State-Trait Anxiety Inventory, trait version; NFCS: Need for Closure Scale;

NFAS: Need for Affect Scale. STAI-T scores range from 20 to 80; NFCS and NFAS use a Likert scale of 6 and 7 points, respectively.

Table 4

Mean values and standard deviations (in parenthesis) of cumulative percentages of correct responses, separately for each image level and emotional valence

Image Level	Neutral	Unpleasant	Pleasant
1	0.07 (0.29)	0.00 (0.00)	0.00 (0.00)
2	3.53 (4.61)	2.74 (4.58)	1.06 (2.79)
3	34.96 (15.32)	31.70 (16.85)	22.67 (16.41)
4	79.71 (15.23)	71.98 (16.74)	66.66 (15.94)
5	94.65 (7.02)	92.95 (7.23)	92.28 (4.95)
6	100.00 (0.00)	100.00 (0.00)	100.00 (0.00)

Table 5

Results of the mixed proportional odds model (behavioral results)

Comparison	Odds Ratio (95% CI)	p-value
Neutral vs. Unpleasant	1.41 (1.21, 1.63)	< .001
Neutral vs. Pleasant	2.01 (1.74, 2.33)	< .001
Unpleasant vs. Pleasant	1.43 (1.23, 1.66)	< .001

Note. An odds ratio larger than 1 (smaller than 1, respectively) implies that the probability of recognition at earlier level is higher (smaller, respectively) for the first versus the second condition in the comparison.

Figure 1A

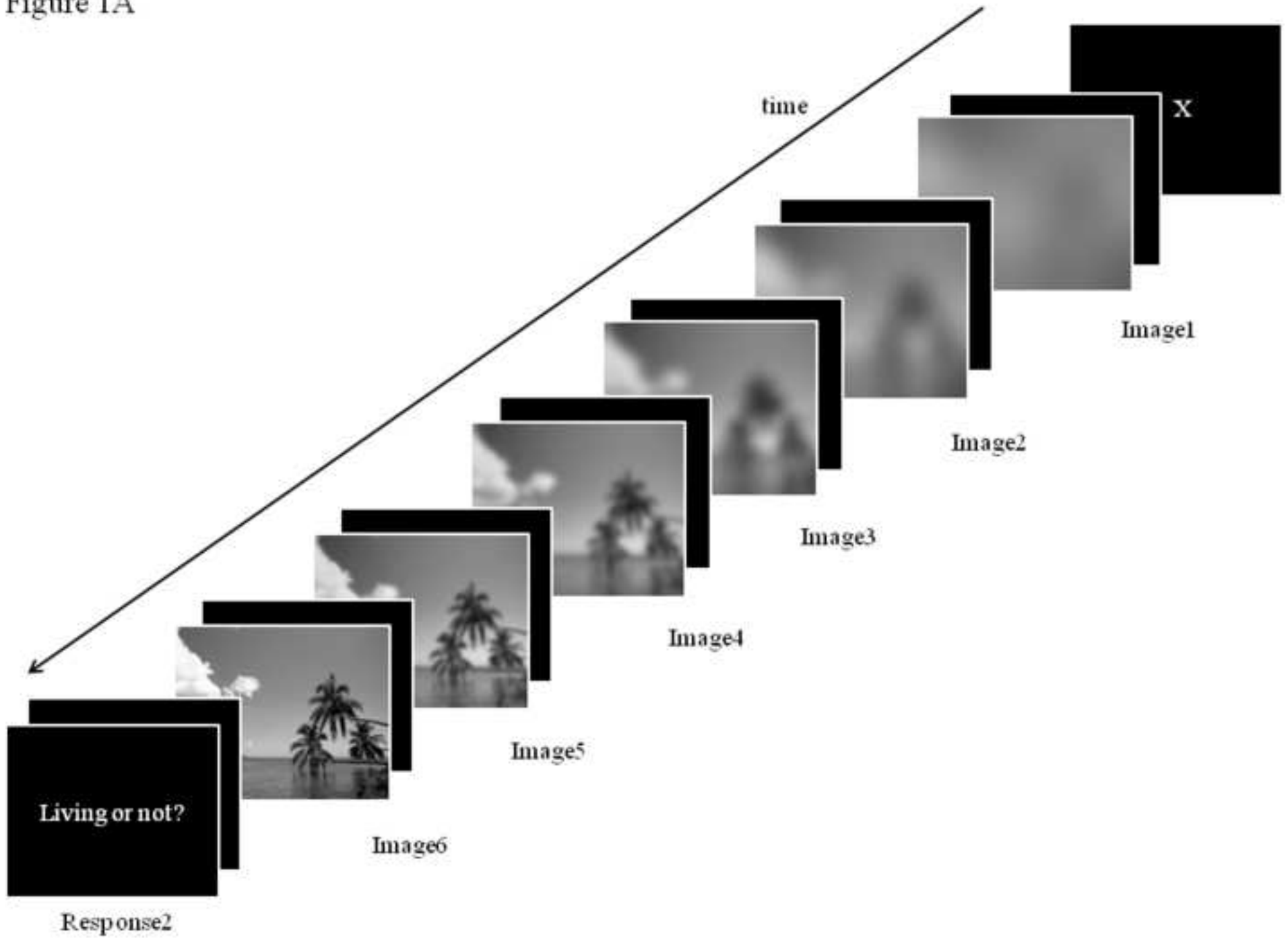


Figure 1B

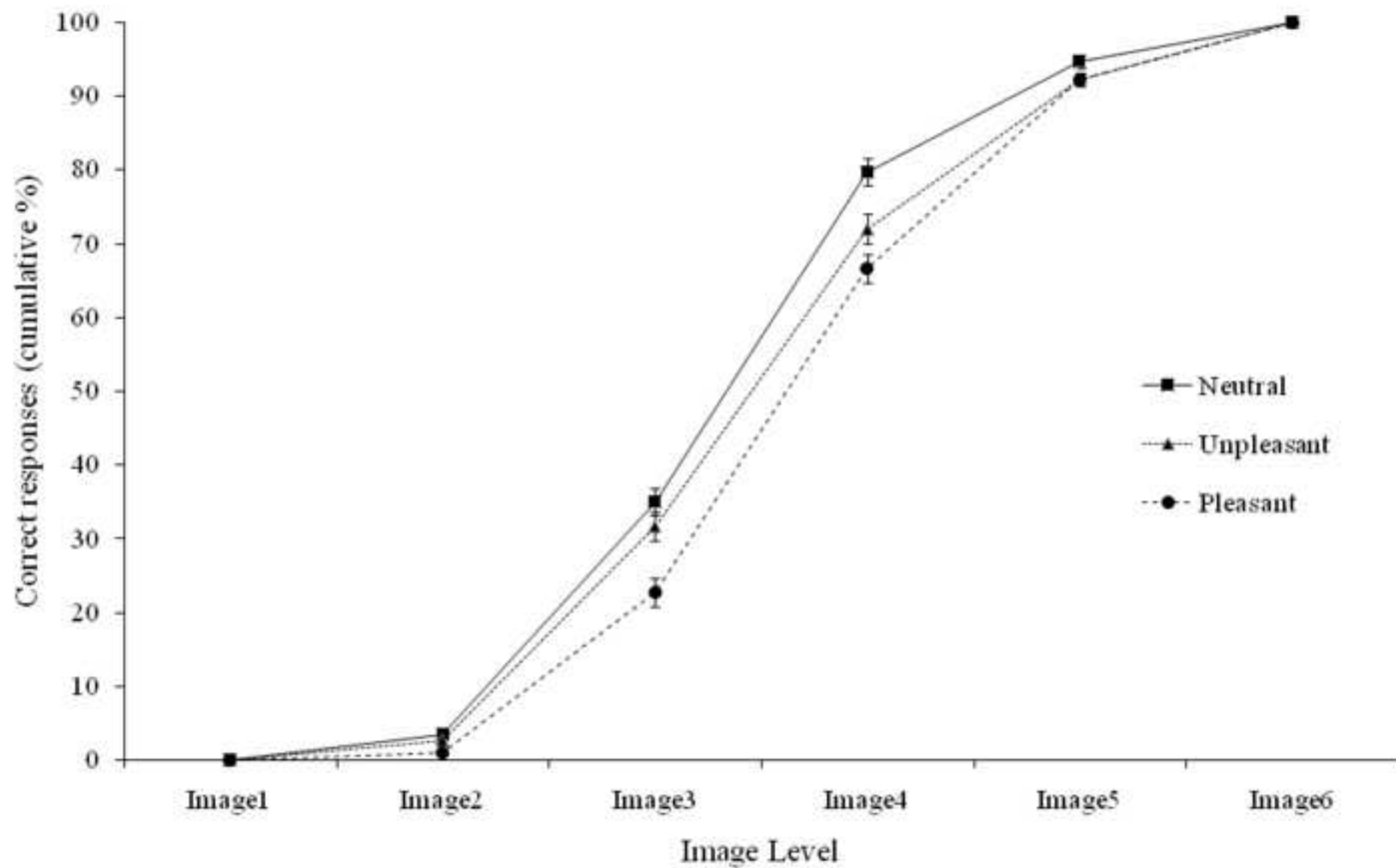


Figure 2A

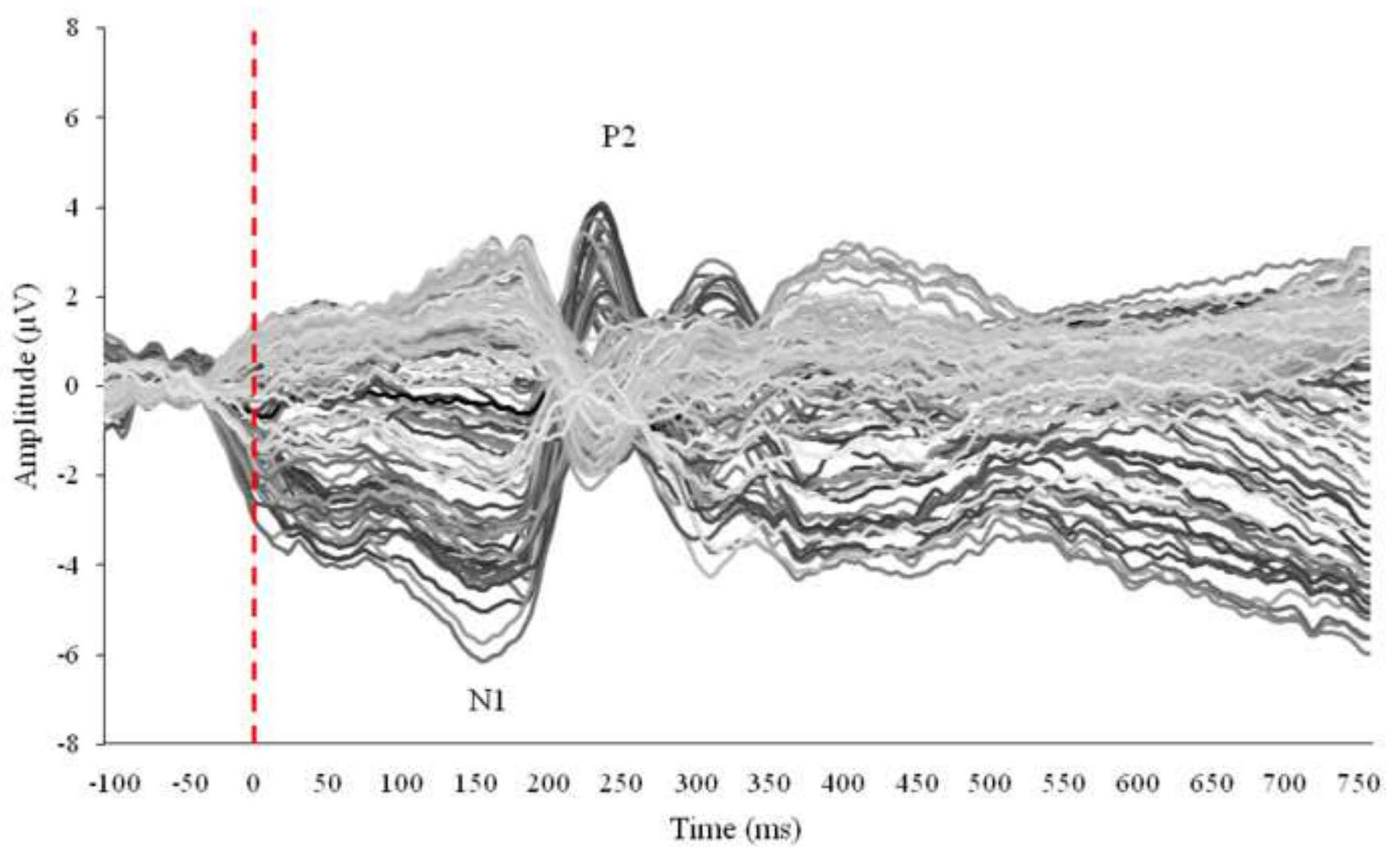


Figure 2B

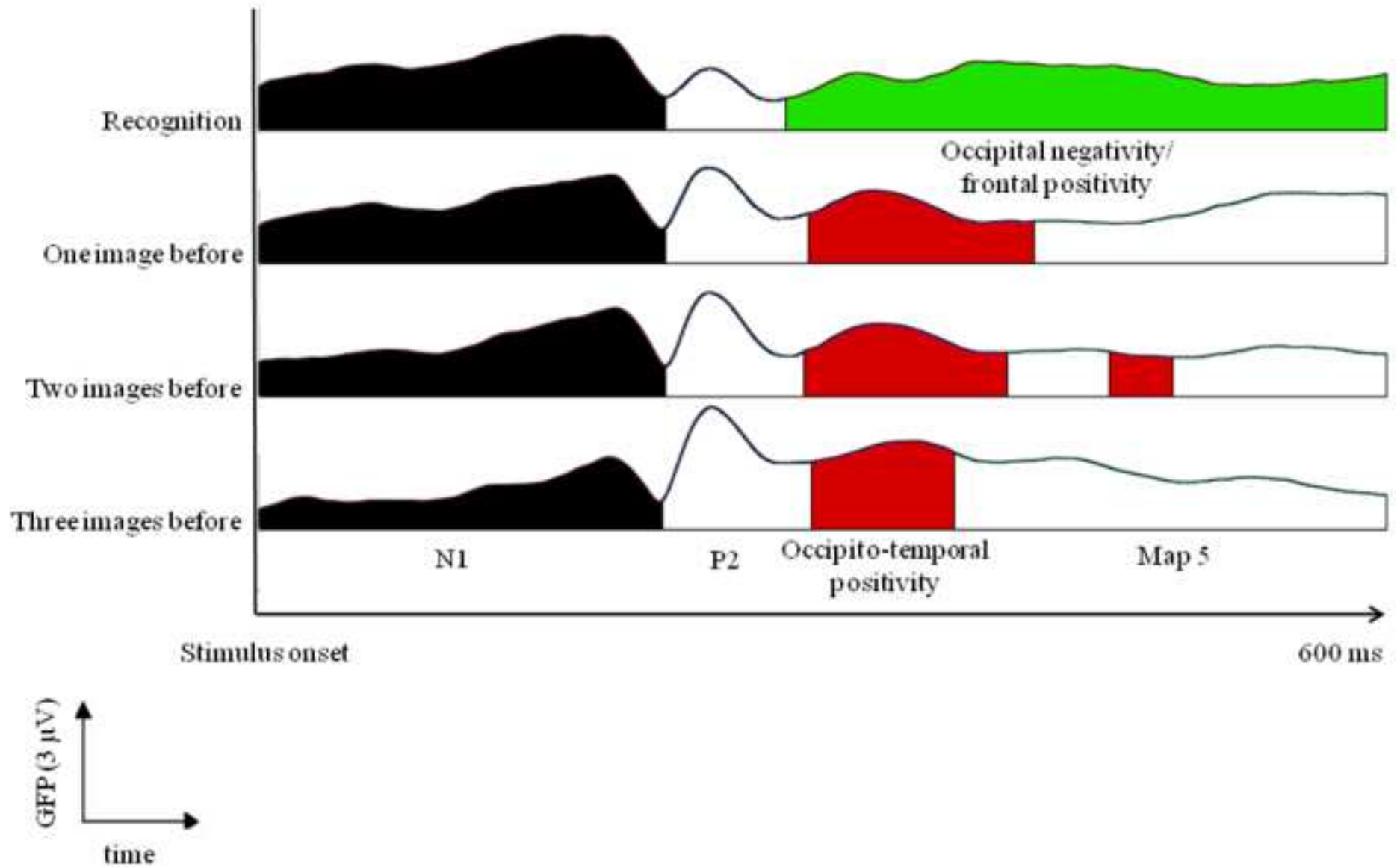


Figure 2C

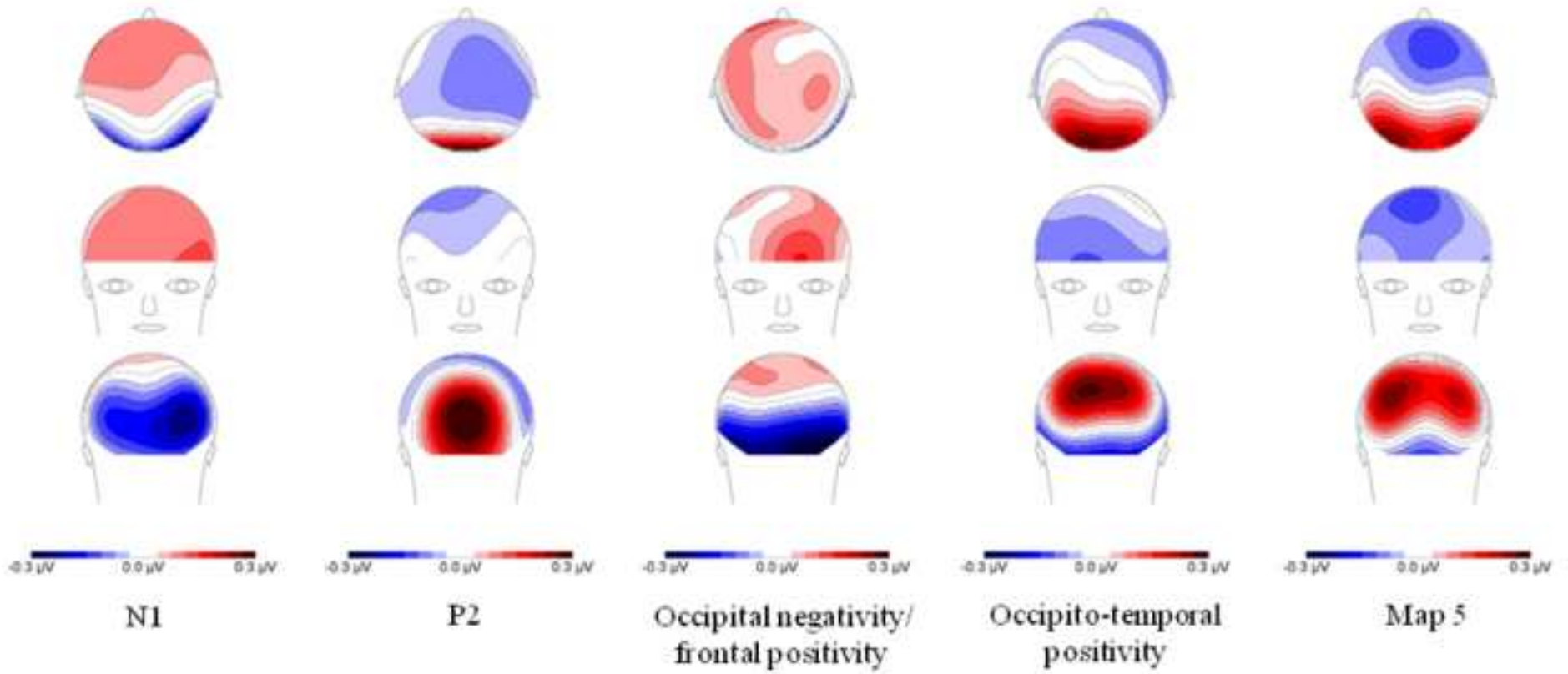


Figure 3A

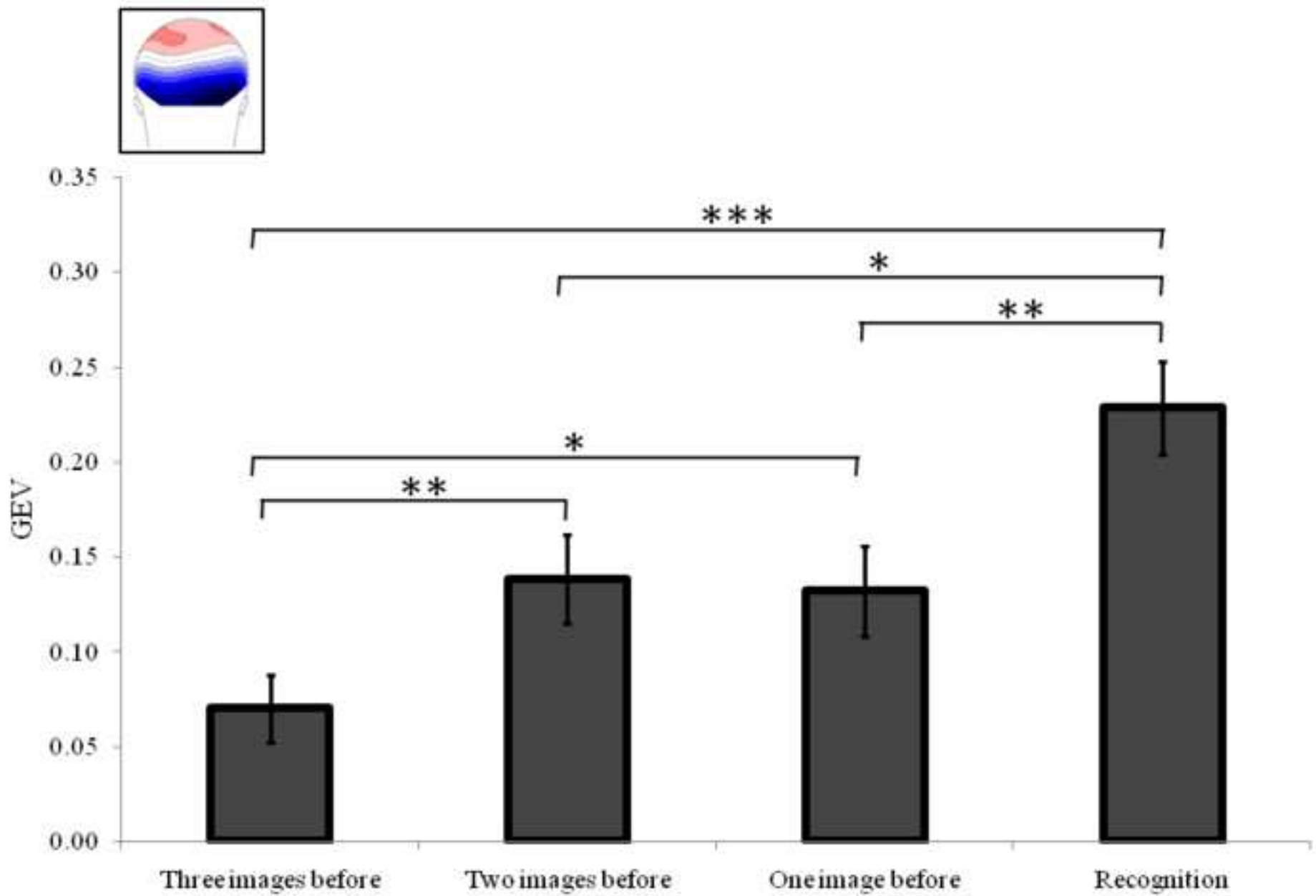


Figure 3B

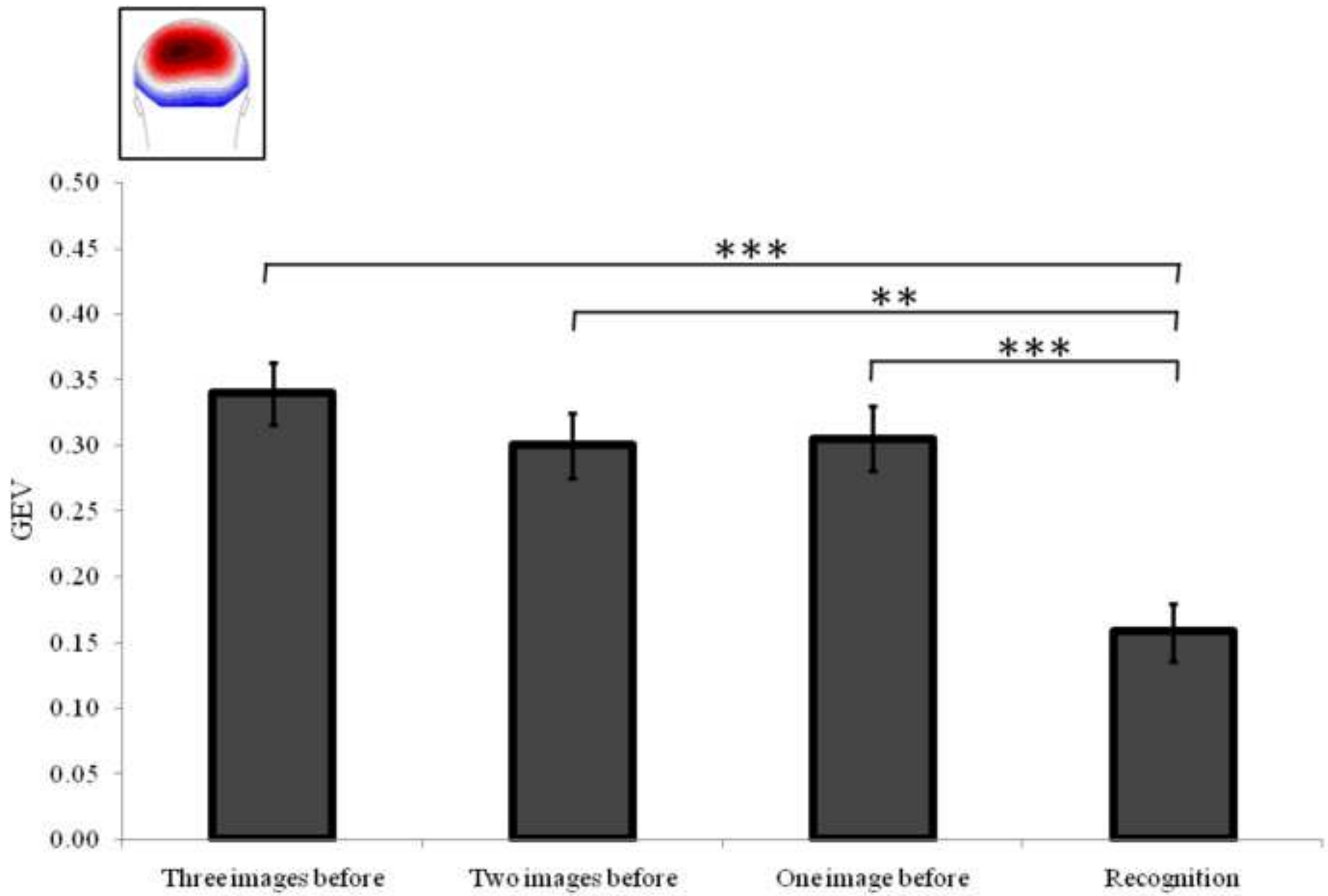


Figure 4A

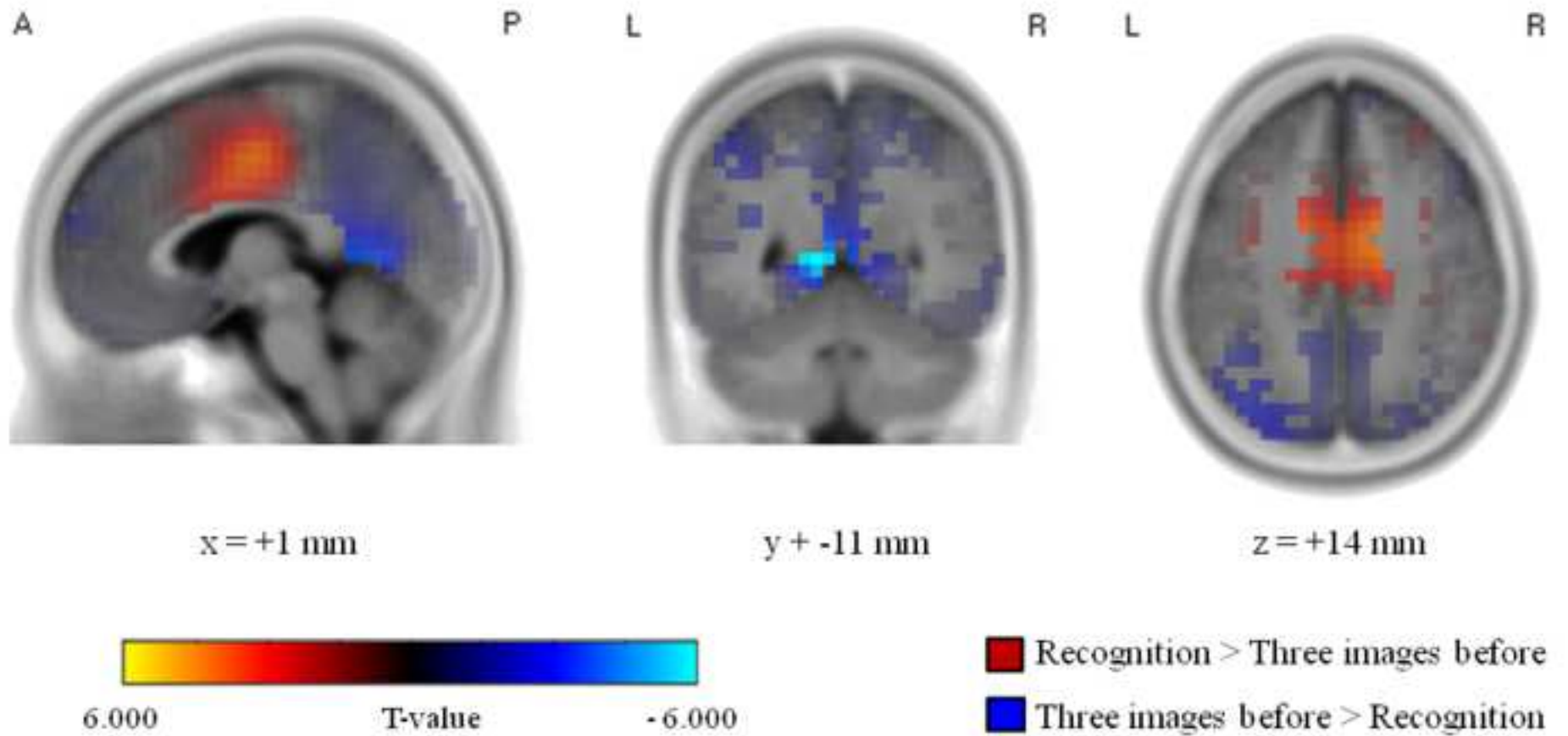


Figure 4B

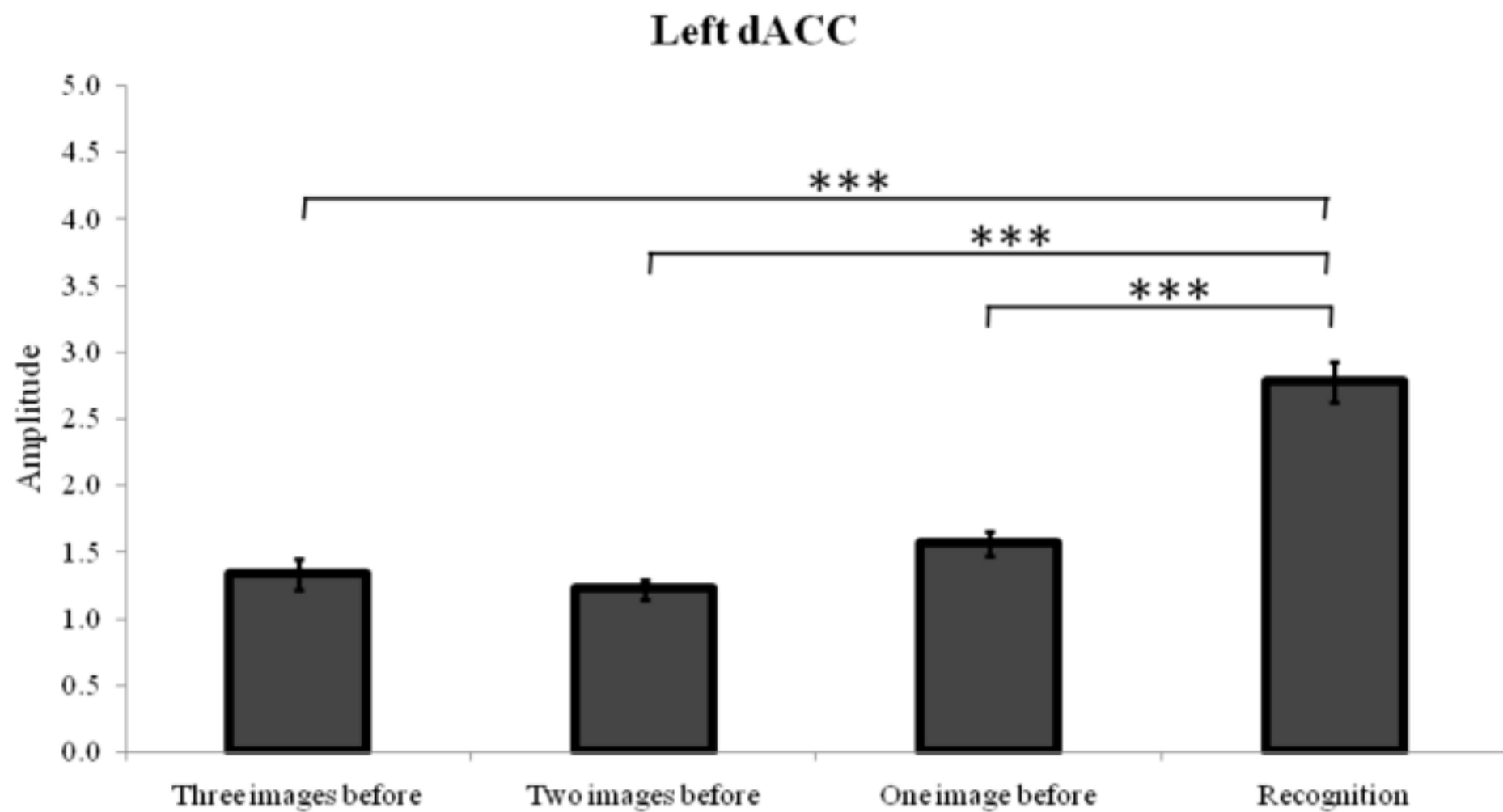


Figure 4C

