Time will tell: Object categorization and emotional engagement during processing of degraded natural scenes

Published Version:
Codispoti, M., Micucci, A., De Cesarei, A. (2021). Time will tell: Object categorization and emotional engagement during processing of degraded natural scenes. PSYCHOPHYSIOLOGY, 58(1), 1-16 [10.1111/psyp.13704].

Availability:
This version is available at: https://hdl.handle.net/11585/784736 since: 2024-05-24

Published:
DOI: http://doi.org/10.1111/psyp.13704

Time will tell: object categorization and emotional engagement during processing of degraded natural scenes

Maurizio Codispoti, Antonia Micucci, & Andrea De Cesarei

Department of Psychology, University of Bologna, Bologna, Italy

Corresponding author: Maurizio Codispoti

Alma Mater Studiorum – University of Bologna

Department of Psychology

Viale Berti Pichat 5, 40127 Bologna

ITALY

e-mail: maurizio.codispoti@unibo.it

Abstract

The aim of the present study was to examine the relationship between object categorization in natural scenes and the engagement of cortico-limbic appetitive and defensive systems (emotional engagement) by manipulating both the bottom-up information and the top-down context. Concerning bottom-up information, we manipulated the computational load by scrambling the phase of the spatial frequency spectrum, and asked participants to classify natural scenes as containing an animal or a person. The role of the top-down context was assessed by comparing an incremental condition, in which pictures were progressively revealed, to a condition in which no probabilistic relationship existed between each stimulus and the following one. In two experiments, the categorization and response to emotional and neutral scenes were similarly modulated by the computational load. The Late Positive Potential (LPP) was affected by the emotional content of the scenes, and by categorization accuracy. When the phase of the spatial frequency spectrum was scrambled by a large amount (>58%), chance categorization resulted, and affective LPP modulation was eliminated. With less degraded scenes, categorization accuracy was higher (.82 in Experiment 1, .86 in Experiment 2) and affective modulation of the LPP was observed at a late window (>800 ms), indicating that it is possible to delay the time of engagement of the motivational systems which are responsible for the LPP affective modulation. The present data strongly support the view that semantic analysis of visual scenes, operationalized here as object categorization, is a necessary condition for emotional engagement at the electrocortical level (LPP).

Introduction

Humans and non-human primates are able to categorize natural scenes rapidly. This extraordinary capacity provides an evolutionary advantage, as our perceptual system serves the adaptive function of categorizing objects as a potential threat or reward in order to act appropriately. Emotional processing is of fundamental importance for survival, and there is a longstanding debate as to whether emotional stimuli are a special class of stimuli, the processing of which is sensitive to the same factors that affect the processing of "neutral" stimuli (Bradley, 2009; Brosch, Pourtois, & Sander, 2010; Harris & Pashler, 2004; Folk, 2015; Öhman, 2005; Pessoa, 2008; Moors, 2007).

Previous studies have systematically examined the impact of natural scene perception on the engagement of cortico-limbic appetitive and defensive systems (hereafter referred to as emotional engagement; Bradley, Codispoti, Cuthbert, & Lang, 2001; Calvo & Nummenmaa, 2007; Ferrari, Mastria, & Codispoti, 2020; Sabatinelli, Keil, Frank, & Lang, 2013; Schupp et al., 2004). Emotional scenes elicit a broad range of autonomic, cortical, behavioral, and subjective changes that reflect the engagement of motivational systems (Lang & Bradley, 2010). From a biphasic motivational view, these responses serve different functions—mobilization for action, attention, and social communication—and reflect the motivational system that is engaged (defensive or appetitive) and its intensity of activation (Bradley, 2009; Lang, Bradley, & Cuthbert, 1997). Consistently, it has been shown that these patterns of physiological reflex reactions vary with evaluative reports of affective arousal and hedonic valence (Bradley, 2000; Cuthbert, Schupp, Bradley, Birbaumer, & Lang, 2000; Lang, Greenwald, Bradley, & Hamm, 1993). Recently, research demonstrated that emotional scenes (pleasant and unpleasant) affect cortical responses even under perceptually challenging conditions in which the stimuli are relatively degraded by reducing their visual angle (3° x 2°; Codispoti & De Cesarei, 2007; De Cesarei & Codispoti, 2006), presenting them in the periphery of the visual field (Calvo, Beltrán, & Fernández-Martín, 2014; De Cesarei, Codispoti, & Schupp, 2009; Keil, Moratti, Sabatinelli, Bradley, & Lang, 2005), or exposing them for a very brief duration (25 ms; Codispoti, Mazzetti, & Bradley, 2009; Ferrari, Codispoti, Cardinale, & Bradley, 2008). However, few studies have directly investigated how emotional engagement varies as a function of both the bottom-up computational load and the top-down setting in which scene categorization takes place.

Visual perception involves bottom–up analysis and top–down modulatory processes, and the interaction of these processes is fundamental for our understanding of natural scenes. Understanding an object or a scene can entail different processes, and result in different perceptual decisions. When vision is unconstrained by specific task demands, basic properties of the environment such as implicit (emotional) relevance or affordances for action can be attained (e.g., Codispoti, Ferrari, De Cesarei, & Cardinale, 2006; Gibson, 1978). However, when a person is carrying out an activity such as categorization or visual search, visual understanding can be guided by templates of the relevant target features (Enns, 2004; Evans & Treisman, 2005). Humans categorize objects accurately even when visual input is impoverished, for example, due to fog, distance, peripheral vision, or poor lighting; under these conditions, even if most fine-grained details are lost, the human visual system is nonetheless very efficient in categorizing objects in natural scenes (e.g., Thorpe, Gegenfurtner, Fabre-Thorpe, & Bülthoff, 2001). On the other hand, stimulus visibility can be hindered by adding noise to an image, or by scrambling the phase structure of a scene; under these conditions, the computational load involved in the perceptual processing increases (Arsenault, Yoonessi, & Baker, 2011; VanRullen, 2011). While it has long been known that perception is an active constructive process, current models of perception suggest that predictive and inferential processes play a fundamental role in modulating object and scene perception (Bar, 2004; Friston & Kiebel, 2009; Schendan & Ganis, 2015). Even for intact scenes, predictions regarding the identity of input objects are crucial, making testing of these predictions a necessary step in the categorization process (Bar, 2007; Friston, 2010; Rao & Ballard, 1999). Accordingly, there is accumulating evidence that the brain quickly makes an initial prediction regarding the 'gist' of natural scenes (Torralba & Oliva 2003), and visual perception is akin to Bayesian Inference (Friston, 2010). This idea, which is thought to find its origins in Helmholtz's notion of "unconscious inference" (unbewusste Schlüsse; Helmholtz, 1867), provides a fundamental theoretical framework for the study of object categorization in natural scenes (see also Gregory, 1980; Friston 2010). Several models suggest that multiple possible interpretations of the input are initially generated during perceptual analysis, and top-down processing guides the choice toward the most likely candidate. These interpretations are then constrained until a choice is made, in light of the previous possible interpretations and of the task at hand (Bar, 2007). At the behavioral level, it has been repeatedly shown that manipulation of the top-down context may facilitate identification (e.g., Sanocki, 1993), but also that in some cases it may lead to the misidentification of stimuli (Bruner & Postman, 1947; Whitson & Galinsky, 2008).

The aim of the present study was to examine the relationship between object categorization in natural scenes and emotional engagement by manipulating both the bottom-up information and the top-down context. Does emotional engagement vary with the computational load involved in categorizing degraded scenes? What is the contribution of top-down processing in the categorization of degraded emotional scenes? Moreover, concerning the specificity of emotional processing, it is possible to hypothesize two possible scenarios. In the first, similar bottom-up and top-down processes underlie the categorization of emotional and neutral scenes. Emotional engagement is expected to decline as the degradation of the stimuli increases, and no emotional engagement is expected to be observed when scenes are not correctly categorized. In the second scenario, categorization of emotional scenes may benefit from specialized processing, which may allow for less pronounced effects of bottom-up or top-down manipulations for emotional compared to non-emotional stimuli.

In the present experiments, the bottom-up computational load was manipulated by scrambling the phase of the spatial frequency spectrum. This manipulation has been shown to be particularly disruptive for the understanding of natural scenes (Wichmann, Braun, & Gegenfurtner, 2006), which relies not only on lower-order visual features such as luminance or contrast, but also on higher-level compositional features (Braun, 2003; Felsen & Dan, 2005). Moreover, scrambling the phase of the spatial frequency spectrum increases categorization difficulty and computational load while keeping lower-order visual features unaltered (Arsenault, et al., 2011; Joubert, Rousselet, Fabre-Thorpe, & Fize, 2009; VanRullen, 2011). On the other hand, the top-down context was manipulated by comparing an incremental condition in which the scenes are progressively revealed (sequential) to a condition in which no probabilistic relationship exists between each stimulus and the following one (mixed). When the scene is initially degraded and then progressively revealed it is likely that the participants develop hypotheses regarding the content of the upcoming scene, and that these previous hypotheses act as a top-down context which modulates the processing of the following picture (Bruner & Potter, 1964; Gollin, 1960). In contrast, when there is no probabilistic relationship between each picture and the following one it can be assumed that the processing will mostly reflect bottom-up stimulus analysis. Following several previous studies, emotional engagement was assessed at the electrocortical level, by measuring the affective modulation (emotional minus neutral) of the Late Positive Potential. Consistent research identifies the late positive potential (LPP) as a reliable cortical marker of emotional processing (Bradley, 2009; Keil et al., 2002; Schupp, Flaisch, Stockburger, Junghöfer, 2006), reflecting both a mandatory engagement of cortico-limbic motivational systems, which continues to take place regardless of stimulus novelty, and enhanced attentional allocation to emotional stimuli (Codispoti, De Cesarei, Biondi & Ferrari, 2016; Ferrari, Bradley, Codispoti, & Lang, 2011; Ferrari, De Cesarei, Mastria, Lugli, Baroni, Nicoletti, & Codispoti, 2016).

Materials and Methods

Participants

A total of 20 participants (14 females; mean age = 21.6, SD = 3.1) took part in the experiment for course credits. We estimated minimal sample size using GPower* (Faul, Erdfelder, Lang, & Buchner, 2007), aiming to determine the number of participants necessary, in a within-participants ANOVA with three conditions (pleasant, neutral, unpleasant), to observe an effect size of at least $\eta^2_p = .07$, with .05 alpha-error probability, 80% power, and a correlation among repeated measures of .6. This analysis yielded 19 as the result. All participants had normal or corrected-to-normal vision, and none of them reported current or past neurological or psychopathological problems. Participants had no previous experience with the materials used in this experiment. The experimental protocol was approved by the Ethical Committee of the Department of Psychology at the University of Bologna.

Stimuli and Equipment

A total of 240 pictures were selected for the present experiment. Half of the pictures represented animals, and half depicted people. For people, pictures varied in affective content, depicting either erotic couples (n = 40), people in a neutral context (n = 40), or mutilated bodies (n = 40). Pictures were selected from various sources, including the International Affective Picture System (IAPS; Lang, Bradley, & Cuthbert, 2015), public domain pictures available on the Internet, and from picture scanning from printed magazines. Pictures subtended a visual angle of 28 (horizontal) by 21 (vertical) degrees.

The pictures were modified in the following way. All pictures were converted to grayscale and equated to the same power spectra, brightness and contrast, using a MATLAB-based toolbox (SHINE; Willenbockel et al., 2010). Then, four phase-scrambled versions of each picture were created using a weighted mean phase (WMP) algorithm (Dakin, Hess, Ledgeway, & Achtman, 2002). This procedure consists of three steps. First, the power and phase of the image spectrum are calculated. Then, the phase spectrum of the original image is combined with a random phase, according to a mixing factor which ranges from 100% (only the random phase information is used) to 0% (only the original phase information is used). In the third and final step, the resulting phase spectrum is recombined with the original spectral power, and a picture is obtained which retains the spectral power of the original image, but with a different phase information. Based on pilot data from three participants who did not take part in the final experiments, it was chosen to use phase scrambling parameters of 80%, 65%, 55%, and 0%. These parameters allowed us to obtain a condition which is not degraded, and which serves as a comparison with existing literature; this condition had 0% phase scrambling and will be henceforth referred to as “Intact pictures”. Moreover, in the pilot phase we observed that, below 80% scrambling, participants appeared to perform at chance; therefore, 80% was chosen as the most degraded point. Finally, we selected 65% and 55% for the remaining two degradation levels in order to be closer to the most degraded rather than to the least degraded level. This procedure was chosen in order to have a sufficient sampling of the rising part of the degradation-accuracy function.

Procedure

The Block and trial procedures are presented in Figure 1. The experiment was divided into two blocks, the order of which was counterbalanced across participants. In one block (mixed block), the order of pictures was pseudorandomized. In neighboring trials of the mixed block, a different picture would always be presented. In designing the mixed block, we wanted to make sure that previous viewing of an intact version of a picture would not modulate the categorization of a following degraded version of the same image; to this end, we organized the mixed block into four subblocks (hence, of 120 pictures each). The presentation order of individual pictures in the four mixed subblocks was the same, so that between two repetitions of the same individual pictures (in different degradation conditions) there were exactly 120 different pictures. Moreover, we assigned each individual picture to either an increasing or a decreasing order. Pictures in increasing order were shown in the most degraded condition (80% phase scrambling) during the first mixed subblock, with 65% phase scrambling during the second mixed subblock, 55% phase scrambling during the third mixed subblock, and in the intact version in the final mixed subblock. Viceversa, pictures in decreasing order were shown in the intact condition (0% phase scrambling) during the first mixed subblock, with 55% phase scrambling during the second mixed subblock, 65% phase scrambling during the third mixed subblock, and in the most degraded version in the final mixed subblock. This strategy allowed us to minimize order effects due to the high number of interposed items (120) between two repetitions of the same picture (which, additionally, was seen in two different degradation conditions), and directly compare the effects of having previously seen a less degraded version of a picture (decreasing condition) to having previously seen a more degraded version of the same picture (increasing condition). To anticipate the results, the increasing or decreasing presentation order did not affect categorization accuracy. In the other block (sequential block), all four versions of the same picture were presented in a row, progressively revealing picture content, from the most degraded to the Intact condition. Pictures were only presented in one of the two blocks, and all pictures were equally often assigned to either block across all participants. In each block, all four versions of the pictures were presented. The two blocks were otherwise identical.

Each trial began with the presentation of a fixation cross, which remained onscreen for 500 ms. Then, a picture was presented and remained onscreen for 1 s. After picture offset a question mark appeared, signaling that a response was required. Participants were instructed to decide whether the picture they had just seen represented a person or an animal by pressing one of two alternative keys (Z or M) on the computer keyboard. Participants were required to respond to all trials, and the association between the response key and the category was counterbalanced across participants. After an intertrial interval of 2 s, the next trial began.

Before the beginning of the experiment, 8 practice trials in the mixed or in the sequential conditions were presented, in order to let participants familiarize themselves with the categorization task. The practice trials did not include any picture which was used during the experiment, and data from the practice trials were not analyzed. Brief breaks were introduced halfway though each block, and between the two blocks.

EEG Recording and Processing

EEG was recorded at a sampling rate of 512 Hz from 256 active sites using an ActiveTwo Biosemi system. An additional sensor was placed below the participant's left eye, to allow for detection of blinks and eye movements. All signals were recorded in single-ended mode. The EEG was referenced in real time to a feedback loop, which comprises a common mode active electrode located near Cz (CMS=common mode sense), and a passive electrode (DRL=driven right leg). A hardware fifth-order low-pass filter with a -3dB attenuation factor at 50 Hz was applied online. Off-line analysis was performed using Emegs (Peyk, De Cesarei, & Junghöfer, 2011). EEG data were initially filtered (0.1 Hz high-pass and 40 Hz low-pass), and eye movements were corrected by means of an automated regressive method (Schlögl et al., 2007). Trials and sensors containing artifactual data were detected through a statistical procedure specifically developed for dense-array EEG (Junghöfer, Elbert, Tucker, &

Rockstroh, 2000). Trials containing a high number of neighboring bad sensors were discarded; for the rest of the trials, sensors containing artifactual data were replaced by interpolating the nearest good sensors. Finally, data were re-referenced to the average of all sensors, and a baseline correction was performed, based on the 100 ms prior to stimulus onset. The percentage of trials that were discarded by the artifact detection procedure was as follows: 80% phase scrambling, 10.64% trials; 65% phase scrambling, 5.42% trials; 55% phase scrambling, 5.17% trials; 0% phase scrambling, 11.17% trials.

Data analysis

The analysis focused on correctly categorized scenes representing people (pleasant, neutral or unpleasant). Electrodes for scoring LPP were decided based on the maximal difference between arousing and neutral scenes in the 400-1200 time interval, with the constraint that the resulting sensor group should be continuous (ie., with no missing sensors in the middle), and symmetrical. The sensor group that was thus obtained is shown in Figure 3.

Results were analyzed through repeated measures ANOVAs with factors Procedure (mixed vs. sequential), Phase scrambling (80%, 65%, 55%, Intact), and Emotional Category (pleasant, neutral, or unpleasant). In order to deal with violations of sphericity, a Huynh-Feldt correction was applied to the degrees of freedom. For all ANOVA effects, we calculated, and report, the partial eta squared (η^2_p), which reflects the proportion of variance that is accounted for by experimental manipulations. Based on previous studies which reported variable latency of the LPP (MacNamara, 2018), we included an additional factor Time (400-800 vs. 800-1200 ms). In behavioral analyses, accuracy in all phase-scrambling conditions was compared to the .50 chance level through a t-test with $\alpha = .05$.

Results

Categorization

Means and standard deviations for categorization accuracy are reported in Table 1, and in Figure 2. Accuracy was initially at chance, and increased as pictures were revealed (.49, .58, .82, and .99, respectively). A significant effect of Phase scrambling was observed on accuracy, $F(3, 57) = 261.98$, $p < .001$, $\eta^2_p = .93$, with more accurate categorization performance as pictures were revealed. Except for the most phase-scrambled (80%) level, all other levels significantly differed from chance, $p_s < .006$. This effect was further qualified by a significant interaction with Emotional Category, $F(6, 114) = 18.59$, $p < .001$, $\eta^2_p = .50$. Following this interaction, accuracy was analyzed for each phase scrambling condition. In the two intermediate (65% and 55%) conditions, a significant effect of Emotional Category was observed, $F_s(2, 38) > 15.50$, $p_s < .001$, $\eta^2_{ps} > .45$, indicating significantly lower accuracy for unpleasant compared to pleasant and neutral scenes in both conditions, $p_s < .007$, and lower accuracy for pleasant compared to neutral pictures in the second most intact (55%) condition, $p = .005$. No effect of Emotional Category was observed in the most degraded (80%) condition and for intact pictures. No significant main effects or interactions involving the factor Procedure were observed. Finally, within the Mixed condition, no difference was observed between pictures which increased in visibility (i.e., began degraded and then were progressively made more visible, with 120 pictures between successive versions of each picture) and pictures which decreased in visibility (i.e., began intact and then were progressively made less visible, with 120 pictures in between), $F(3, 57) = 1.528$, $p = .221$, $\eta^2_p = .074$.

LPP

The effects of degradation and emotional category on the LPP are reported in Figure 3. A significant three-way interaction between Time, Phase scrambling, and Emotional Category was observed, $F(6, 114) = 5.08$, $p < .001$, $\eta^2_p = .21$. When analyzing each phase-scrambling level separately, significant interactions of Time and Emotional Category were only observed for the second most intact (55%) phase scrambling condition and intact pictures $F_s(2, 38) > 9.49$, $p_s < .002$, $\eta^2_{ps} > .33$ (see Figures 3 and 4). In the 55% phase scrambling condition, a significant effect of Emotional Category was observed in the 800-1200 time interval, $F(2, 38) = 5.27$, $p = .010$, $\eta^2_p = .22$, indicating a significantly more positive amplitude of the LPP for pleasant compared with neutral or unpleasant scenes, $F_s(1, 19) > 5.38$, $p_s < .032$, $\eta^2_{ps} > .22$; this effect was not observed in the 400-800 time interval, $p = .173$. For intact pictures, significant effects of Emotional Category were observed in both time intervals, $F_s(2, 38) > 30.83$, $p_s < .001$, $\eta^2_{ps} > .62$, with arousing (both pleasant and unpleasant) contents eliciting a significantly more pronounced LPP compared to those that were neutral, $p_s < .001$.

No significant interactions simultaneously involving the Procedure and Emotional Category factors were observed. A significant interaction of Procedure and Phase scrambling was observed, $F(3, 57) = 5.68$, $p = .004$, $\eta^2_p = .23$, indicating that the LPP amplitude was more positively pronounced for the sequential compared to the mixed condition for intact scenes, $F(1, 19) = 15.96$, $p = .001$, $\eta^2_p = .46$, but not for degraded pictures, $p_s > .396$. Finally, a significant interaction of Time and Phase scrambling was observed, $F(3, 57) = 28.58$, $p < .001$, $\eta^2_p = .60$, indicating that in the two intermediate (65% and 55%) conditions the amplitude of the LPP increased significantly over time, $F_s(1, 19) > 15.22$, $p_s < .001$, $\eta^2_{ps} > .45$, while for intact scenes a more positive ERP amplitude was observed in the 400-800 compared to the 800-1200 ms time interval, $F(1, 19) = 13.94$, $p = .001$, $\eta^2_p = .42$.

Discussion of Experiment 1

In Experiment 1, natural scenes were degraded and categorization was assessed in a two-alternative forced choice task. Accuracy increased as scenes were revealed, and a more positive LPP amplitude was observed for intact arousing, compared to neutral, scenes (Schupp et al., 2006). Concerning degraded scenes however, when categorization accuracy was low (below .60) no affective modulation of the LPP was observed. When pictures were more identifiable (.83 accuracy in the 55% phase scrambling condition), a significant affective ERP modulation with direction and topography consistent with the LPP was observed, beginning 800 ms after picture onset.

In Experiment 1 the top-down setting did not have any noticeable effect on the affective modulation of the LPP, and neither a facilitation nor an interference was observed in categorization performance. One likely explanation is that little information is collected from the two most degraded picture versions (i.e., 80% and 65% phase scrambling), eliminating the possibility for prior knowledge to modulate further processing. To test this possibility, we conducted a second experiment, in which pictures were less degraded. According to this reasoning, we expected a higher accuracy in the most degraded conditions to allow for more pronounced top-down effects.

Experiment 2

Materials and Methods

Participants

As in Experiment 1, a total of 20 participants (12 females; M = 28.5, SD = 5.4) participated in the experiment as volunteers. Participants had normal or corrected-to-normal vision, and none of them reported current or past neurological or psychopathological problems. Participants had no previous experience with the materials used in this experiment. The experimental protocol was approved by the Ethical Committee of the Department of Psychology at the University of Bologna.

Scrambling parameters and procedure

Picture manipulation was identical to that of Experiment 1 and included conversion to grayscale and balancing of the power spectra, brightness, and contrast. The only difference compared to Experiment 1 was in the parameters for phase scrambling, which were chosen to be overall easier to categorize compared to Study 1, and were 58%, 55%, 50%, and 0% (Intact). Importantly, the 55% phase scrambling level replicates the condition in Experiment 1 in which affective LPP modulation was delayed. In Experiment 2 the duration of the fixation cross varied between 500 and 1000 ms, to avoid participants anticipating stimulus onset, and this anticipation being reflected in the ERP baseline.

Analysis

Collection and analysis of behavioral responses and ERP data were similar to those of Experiment 1. The LPP was scored using the same region and temporal windows of interest used in Experiment 1. The percentage of trials discarded by the artifact detection procedure was as follows: 58% phase scrambling, 9.28% trials; 55% phase scrambling, 8.52% trials; 50% phase scrambling, 9.12% trials; 0% phase scrambling, 8.66% trials.

Results

Categorization

Categorization accuracy is reported in Table 2. Similarly to Experiment 1, a significant effect of Phase scrambling was observed on accuracy, $F(3, 57) = 164.01$, $p < .001$, $\eta^2_p = .90$. Accuracy was initially .76, and increased as scenes were revealed (.86, .96, and 1.00, respectively). At all phase-scrambling levels, accuracy was significantly above chance, $p_s < .001$. This effect was further qualified by a significant interaction with Emotional Category, $F(6, 114) = 36.40$, $p < .001$, $\eta^2_p = .66$. Analyzing the effects of categorization at each phase-scrambling level, significant effects of Emotional Category were observed in the three degraded (58% to 50%) phase-scrambling conditions, $F_s(2, 38) > 33.23$, $p_s < .001$, $\eta^2_{ps} > .64$, with less accurate categorization for unpleasant compared to all other categories in all degraded levels, $p < .001$, and for pleasant compared to neutral pictures in the two most degraded (58% and 55%) conditions, $p < .003$. No effect of Emotional Category was observed for intact pictures ($p = .35$).

A significant interaction was observed between Phase scrambling and Procedure, $F(3, 57) = 6.23$, $p = .003$, $\eta^2_p = .25$. Following this interaction, significant effects of Procedure were observed in the two intermediate (55% and 50%) phase-scrambling levels, $p_s < .007$, with slightly but significantly higher accuracy in the mixed compared to the sequential condition (.89 vs. .83 for the 55% phase-scrambling level, .97 vs. .95 for the 50% phase scrambling condition). Within the Mixed condition, no difference was observed between pictures that increased in visibility and pictures that decreased in visibility, $F(3, 57) = 1.948$, $p = .167$, $\eta^2_p = .093$.

LPP

The effects of picture degradation on the affective modulation of the LPP are shown in Figure 5. Similarly to Experiment 1, a significant interaction of Time, Phase scrambling, and Emotional Category was observed, $F(6, 114) = 12.58$, $p < .001$, $\eta^2_p = .40$. When decomposing this interaction at each level of phase scrambling, significant interactions of Time and Emotional Category were observed for the two intermediate (55% and 50%) phase scrambling conditions and for intact pictures, $F_s(2, 38) > 4.55$, $p_s < .017$, $\eta^2_{ps} > .19$. A significant effect of Emotional Category was observed in both the 400-800 ms and 800-1200 ms time intervals in the 55% phase-scrambling condition $F_s(2, 38) > 3.48$, $p < .048$, $\eta^2_p > .16$ as well as in the 50% phase scrambling condition $F_s(2, 38) > 19.14$, $p_s < .001$, $\eta^2_{ps} > .50$. In both conditions, while pleasant scenes elicited a significantly more positive LPP compared to neutral scenes in both time intervals (400-800, $F_s(1, 19) > 4.76$, $p_s < .042$, $\eta^2_{ps} > .20$; 800-1200, $F_s(1, 19) > 12.28$, $p_s < .002$, $\eta^2_{ps} > .39$), the difference between unpleasant and neutral scenes only reached significance in the 800-1200 ms window, $F(1, 19) = 11.02$, $p_s < .004$, $\eta^2_{ps} = .37$, and $F(1, 19) = 30.95$, $p < .001$, $\eta^2_p = .62$ for the 55% and 50% phase scrambling condition respectively[1] (Figures 5 and 6). Concerning intact scenes, significant effects of Emotional Category were observed at both time intervals, $F_s(2, 38) > 47.61$, $p_s < .001$, $\eta^2_{ps} > .72$, with pleasant and unpleasant scenes eliciting a more positive LPP compared to neutral pictures, $p_s < .001$.

A significant interaction of Time, Procedure, and Phase scrambling was observed, $F(3, 57) = 4.77$, $p = .012$, $\eta^2_p = .20$. When this interaction was decomposed for each phase scrambling condition, a significant interaction of Time and Procedure was observed in the second most intact (50%) phase scrambling condition and for intact pictures, $F_s(1, 19) > 17.66$, $p_s < .001$, $\eta^2_{ps} > .48$. In both conditions, no effects of Procedure were observed in the 400-800 time interval, ps > .575, while a more positive LPP was observed in the sequential, compared to the mixed, condition, $F_s(1, 19) > 6.10$, $p_s < .023$, $\eta^2_{ps} > .24$. Although a significant interaction of the Procedure and Emotional Category factors was observed, $F(2, 38) = 3.26$, $p = .049$, $\eta^2_p = .15$, the effect of Procedure did not reach significance for any emotional category, and no other significant interactions simultaneously involving the Procedure and Emotional Category factors were observed.

A significant interaction of Time and Emotional Category was also observed, $F(2, 38) = 9.17$, $p = .001$, $\eta^2_p = .33$, descriptively indicating that the positivity of the LPP for unpleasant scenes increased over time; however, the effects of Time did not reach significance for any emotional category ($p_s > .197$).

General Discussion

The present study aimed to examine the relationship between object categorization in natural scenes and emotional engagement by manipulating both the bottom-up information and the top-down context. In both experiments, the latency and amplitude of the LPP affective modulation varied with stimulus degradation (phase scrambling; see Figure 7) and LPP amplitude was modulated by affective picture content only when accuracy was high; moreover, no effects of the top-down context were observed. In sum, the present data strongly support the possibility that the processing of emotional and neutral scenes is similarly modulated by computational load.

Both experiments indicated that it is possible to delay the time of engagement of the motivational systems which are responsible for the LPP affective modulation. Increasing the computational load necessary for object categorization eliminated the affective modulation of cortical activity up to a late window beginning at 800 ms. Only at this latency, for stimuli which were categorized with about .82 accuracy (55% phase-scrambling condition in both studies), was a late positive ERP modulation observed with the same direction and topography as the LPP. When stimuli were made easier to identify, this modulation became more pronounced in amplitude and earlier in latency. Importantly, however, in all observed time ranges ERP affective modulation was only observed when categorization accuracy was remarkably high, supporting the view that semantic processing, operationalized here as object categorization, is a necessary condition for emotional engagement. The disproportion between the degradation level needed to achieve an above-chance categorization, or a significant affective categorization, could be related to the timing of the present paradigm, as well as to the nature of the categorization task. In terms of timing, the categorization task was completed after more than a second from scene onset. In this time interval the processing of the scene could be accomplished, while the ERP affective modulation occurred in response to stimulus onset (about a second earlier). In addition, another critical difference is the nature of the processing involved in the two measures (affective modulation of the LPP and accuracy in the animal/people categorization task). The accuracy in the categorization task reflects the specific process involved, in which participants could rely on the templates of the two possible targets (e.g., features of animals and people) in the categorization of the incoming stimulus. Knowing that only two possible targets are expected facilitates the processing because the stimulus to detect is primed, and the matching of a top-down template with bottom-up information can maximize the efficiency of categorization (Enns, 2004). In contrast, in the case of categorization of the emotional content, reflected in the affective modulation of the LPP, participants were not asked to categorize the emotional content, and given the varied picture content, a top-down template that guides recognition is less likely to be efficient.

The findings of both experiments suggest that semantic understanding of objects in scenes is a necessary condition for the affective modulation of the LPP. Specifically, no affective modulation of the LPP was observed when categorization accuracy was at chance or low, and affective modulation of the LPP was first observed only when categorization performance was high (above .82), while the opposite (affective modulation in the absence of accurate categorization) was never observed. Similar findings have been observed in previous ERP studies in which stimulus degradation was manipulated using different techniques, such as backward visual masking (Codispoti et al., 2009; Grassini, Holm, Railo, & Koivisto, 2016), spatial frequency filtering (high and low spatial frequencies; De Cesarei & Codispoti, 2011), and addition of visual noise (Schupp et al., 2008). However, it should be noted that previous ERP studies used subjective reports (or scene gist) to assess stimulus understanding, and did not rely on a direct (objective) measure of accuracy in scene categorization. Moreover, in a series of experiments using saccade latency, Nummenmaa and Calvo observed findings similar to those described in the present study (Calvo & Nummenmaa, 2007; Nummenmaa, Hyönä, & Calvo, 2010). For example, Nummenmaa and colleagues (2010) asked participants to view paired emotional and neutral pictures (involving humans or animals) briefly presented in extrafoveal vision, and instructed them to categorize the targets by executing a saccade toward the location of a predefined target content. The semantic task involved saccading toward a scene containing an animal, and the emotional task involved saccading toward an unpleasant or pleasant picture. Results indicated that semantic categorization was faster than affective categorization. Consistently, when pictures were foveally presented scenes, and the time available for visual processing was varied using a backward masking procedure, exposure threshold for accurate categorization was lower for semantic information than for affective information. Altogether, these previous studies, and our findings, seem to suggest that affective evaluation and the engagement of corticolimbic systems occur only after semantic categorization of the visual scene had taken place. Consistently, several studies failed to detect evidence of emotion effects (emotional minus neutral) when stimulus recognition was impaired using backward masking or continuous flash suppression paradigms (Hedger, Adams, & Garner, 2015; Peira, Golkar, Öhman, Anders, & Wiens, 2012; Pessoa, 2005), and recent meta-analysis has demonstrated that stimulus visibility moderates emotion effects, which are elicited only after the stimulus is consciously recognized (Hedger, Garner, & Adams, 2016; Lähteenmäki, Hyönä, Koivisto, & Nummenmaa, 2015).

In the present experiments, the affective modulation of cortical changes was only observed when categorization accuracy was remarkably high, while no modulation was found when the accuracy was below .82. Since phase scrambling has the remarkable advantage (compared, for instance, to spatial frequency filtering) of leaving lower order visual features constant (Arsenault et al., 2011; Joubert et al., 2009), the absence of any emotional effect in these conditions clearly demonstrated that the affective modulation of the LPP for intact scenes is not due to physical properties of the images (i.e., contrast or spatial frequency amplitude), but relies on the semantic understanding of picture content. This finding is consistent with previous studies showing that the affective modulation of the LPP reflects emotional engagement, and is not due to perceptual factors, such as spatial frequency content, complexity, or color (Bradley, Hamby, Löw, & Lang, 2007; De Cesarei & Codispoti, 2011; Codispoti, De Cesarei & Ferrari, 2012).

In terms of top-down modulation, the impact of a predictive context was examined by presenting scenes in progressively less degraded versions, and different levels of difficulty were chosen for Experiments 1 and 2. In the sequential condition, once a picture was categorized in a single step, it served as a perceptual and motivational context for the categorization of the following picture. No advantage was observed in the processing of emotional compared to non-emotional scenes, either in terms of categorization or of emotional engagement. In Experiment 1, the top-down context did not have any noticeable effect in terms of overall performance. However, since the first degradation levels did not attain high accuracy, suggesting that little information could be collected in these initial levels, it is not surprising that no effects were observed in the analysis of subsequent picture versions. Basically, effects of a motivational or perceptual top-down context are expected when enough picture content has been attained, and more pronounced effects of procedure are expected when categorization accuracy is higher. Therefore, we conducted Experiment 2 in which we used less degraded scenes in order to observe top-down effects in the sequential, compared to the mixed, condition. A slightly but significantly worse categorization performance was found in the sequential, compared to the mixed, condition as the picture content was progressively revealed. Moreover, similar affective modulation of the LPP was observed in both conditions. These findings suggest that the affective modulation of the LPP is not sensitive to top-down contextual factors even when processing degraded stimuli, but is primarily driven by the recognition of the emotional content in the current stimulus. Consistent with this finding, previous studies by Schupp and colleagues found that the ERPs elicited by emotional pictures were not affected by the hedonic context provided by preceding images, even when several images (60) of the same valence were presented (Flaisch, Junghöfer, Bradley, Schupp, & Lang, 2008; Schupp, Schmälzle, Flaisch, Weike, & Hamm, 2012). Similarly, other studies on repeated stimulus exposure indicated that the emotional modulation of the LPP is mostly driven by the information depicted in the current stimulus with little space for top-down factors (Codispoti, Ferrari, Bradley, 2007; Codispoti et al., 2016; Ferrari, Bradley, Codispoti, Karlsson, Lang, 2013; Micucci, Ferrari, De Cesarei, & Codispoti, 2020).

In terms of categorization accuracy, we observed a small but significant detrimental effect of the sequential, compared to the mixed condition, for the intermediate degradation levels in Experiment 2. These findings are consistent with previous demonstrations of interference, and/or performance reduction, due to erroneous expectations (hypotheses) or choices. An interference effect on visual image identification prompted by previous exposures to its partial features was first demonstrated by Galloway (1946; but see also Wyatt & Campbell , 1951) and is called the “perceptual interference effect” (Luo & Snodgrass 1994). This phenomenon was replicated by Bruner & Potter (1964), and later systematically examined by Snodgrass and colleagues (Snodgrass & Hirschman, 1991; Luo &

Snodgrass,1994). In Bruner & Potter's seminal study (1964), pictures of everyday objects were shown in a descending degradation sequence, beginning from a very, moderately, or slightly blurry initial version, and progressing to a final intact version. Participants had to understand picture content, and their performance declined depending on the initial blur level, suggesting interference produced by previous hypotheses. The inhibitory effect of early exposures to a visual image's partial features on its subsequent identification was later replicated by Snodgrass and colleagues, who developed the *competitive activation model* (Luo & Snodgrass,1994; Snodgrass & Hirschman, 1991). According to this model, during cue presentations in the descending degradation sequence, participants develop erroneous (and competitive) hypotheses about the stimulus that interfere with its correct perception. In a simulation, Snodgrass & Hirschman (1991) showed that the interference effect can be produced by adding transient activations in perceptual structures that are generated by cue presentations in the descending procedure. More recently, even simple forced-choice decisions regarding perceptual stimuli have been shown to elicit poorer performance due to choice history biases (Abrahamyana, Silva, Dakin, Carandini, & Gardner, 2016). Overall, these interference effects might be related to a visual confirmation bias; this bias connotes "the seeking or interpreting of evidence in ways that are partial to existing beliefs, expectations, or a hypothesis in hand" (Nickerson, 1998). The perceptual interference effect and the choice history bias reported in these previous studies might explain the reduced accuracy in the sequential condition, compared to the mixed one, observed in Experiment 2.

Conclusion

Here we investigated the relationship between object categorization in real-world scenes and emotional engagement by manipulating both the bottom-up information and the top-down context. The results clearly indicate that emotional engagement, indexed by the affective modulation of the late positive potential, is primarily driven by the emotional content in the current stimulus, and is not sensitive to top-down contextual factors even when degraded stimuli are being processed. Moreover, the affective modulation of the LPP was observed when categorization accuracy was remarkably high, supporting the view that object categorization is a necessary condition for the engagement of cortico-limbic appetitive and defensive systems in the perception of real-world scenes.

Acknowledgements

The authors thank all participants for having taken part in the study.

References

Abrahamyan, A., Silva, L. L., Dakin, S. C., Carandini, M., & Gardner, J. L. (2016). Adaptable history biases in human perceptual decisions. *Proceedings of the National Academy of Sciences* USA, 113, E3548–E3557.

Arsenault, E. , Yoonessi, A., and Baker Jr., C. (2011). Higher order texture statistics impair contrast boundary segmentation. *Journal of Vision* 11, 1-15. https://doi.org/10.1167/11.10.14

Bar, M. (2007). The proactive brain: using analogies and associations to generate predictions. *Trends in Cognitive Sciences*, 11(7), 280–9. https://doi.org/10.1016/j.tics.2007.05.005

Bar, M. (2004). Visual objects in context. *Nature Reviews Neuroscience*, 5, 617–629. https://doi.org/10.1038/nrn1476

Bradley, M. M., Hamby, S., Löw, A., & Lang, P. J. (2007). Brain potentials in perception: Picture complexity and emotional arousal. *Psychophysiology*, 44, 364–373. https://doi.org/10.1111/j.1469-8986.2007.00520.x

Bradley, M. M. (2009). Natural selective attention: Orienting and emotion. *Psychophysiology*, 46, 1–11. https://doi.org/10.1111/j.1469-8986.2008.00702.x

Bradley, M. M., Codispoti, M., Cuthbert, B. N., & Lang, P. J. (2001). Emotion and motivation I: Defensive and appetitive reactions in picture processing. *Emotion*, 1(3), 276–298. https://doi.org/10.1037/1528-3542.1.3.276

Bradley, M.M. (2000). Emotion and Motivation. In: Cacioppo, Tassinary, & Berntson (Eds): Handbook of Psychophysiology, 2nd Ed. Cambridge University Press.

Braun, J. (2003). Natural scenes upset the visual applecart. *Trends in Cognitive Sciences*, 7, 7-9. https://doi.org/10.1016/S1364-6613(02)00008-6.

Brosch, T., Pourtois, G., & Sander, D. (2010). The perception and categorization of emotional stimuli: A review. *Cognition & Emotion*, 24, 377–400. https://doi.org/10.1080/02699930902975754

Bruner, J. S., & Postman, L. (1947). Tension and tension release as organizing factors in perception. *Journal of Personality*, 15, 300–308. https://doi.org/10.1111/j.1467-6494.1947.tb01070.x

Bruner, J. S., & Potter, M. C. (1964). Interference in visual recognition. *Science*, 144, 424–425. https://doi.org/10.1126/science.144.3617.424

Calvo, M. G., & Nummenmaa, L. (2007). Processing of unattended emotional visual scenes. *Journal of Experimental Psychology: General,* 136, 347–369.

Calvo, M. G., Beltrán, D., & Fernández-Martín, A. (2014). Early processing of happy facial expressions in peripheral vision: Neurophysiological evidence. *Biological Psychology,* 100, 60–70. doi:10.1016/j. biopsycho.2014.05.007

Codispoti, M., Ferrari, V., De Cesarei, A., Cardinale, R. (2006). Implicit and explicit categorization of natural scenes. *Progress in Brain Research, 156,* 57–69. https://doi.org/10.1016/S0079-6123(06)56003-0

Codispoti, M., & De Cesarei, A. (2007). Arousal and attention: Picture size and emotional reactions. *Psychophysiology,* 44(5), 680–686. https://doi.org/10.1111/j.1469-8986.2007.00545.x

Codispoti, M., Ferrari, V., & Bradley, M. M. (2007). Repetition and event-related potentials: distinguishing early and late processes in affective picture perception. *Journal of Cognitive Neuroscience*, 19, 577–86. https://doi.org/10.1162/jocn.2007.19.4.577

Codispoti, M., Mazzetti, M., & Bradley, M. M. (2009). Unmasking emotion: exposure duration and emotional engagement. *Psychophysiology*, 46, 731–8. https://doi.org/10.1111/j.1469-8986.2009.00804.x

Codispoti, M., De Cesarei, A., & Ferrari, V. (2012). The influence of color on emotional perception of natural scenes. *Psychophysiology*, 49, 11–16. https://doi.org/10.1111/j.1469-8986.2011.01284.x

Codispoti, M., De Cesarei, A., Biondi, S., & Ferrari, V. (2016). The fate of unattended stimuli and emotional habituation: Behavioral interference and cortical changes. *Cognitive Affective and Behavioral Neuroscience, 16,* 1063-1073. https://doi.org/10.3758/s13415-016-0453-0

Cuthbert, B. N., Schupp, H. T., Bradley, M. M., Birbaumer, N., & Lang, P. J. (2000). Brain potentials in affective picture processing: Covariation with autonomic arousal and affective report. *Biological Psychology*, 52, 95–111. https://doi.org/10.1016/S0301-0511(99)00044-7

Dakin, S.C., Hess, R.F., Ledgeway, T., and Achtman, R.L. (2002). What causes non-monotonic tuning of fMRI response to noisy images? *Current Biology* 12, 476–477. https://doi.org/10.1016/S0960-9822(02)00960-0

De Cesarei, A., & Codispoti, M. (2006). When does size not matter? Effects of stimulus size on affective modulation. *Psychophysiology*, 43, 207–15. https://doi.org/10.1111/j.1469-8986.2006.00392.x

De Cesarei, A., & Codispoti, M. (2011). Scene Identification and Emotional Response: Which Spatial Frequencies Are Critical? *Journal of Neuroscience*, 31, 17052–17057. https://doi.org/10.1523/JNEUROSCI.3745-11.2011

De Cesarei, A., Codispoti, M., & Schupp, H. T. (2009). Peripheral vision and preferential emotion processing: *Neuroreport*, 20, 1439–1443. https://doi.org/10.1097/WNR.0b013e3283317d3e

Enns, J. T. (2004). Object substitution and its relation to other forms of visual masking. Vision Research, 44, 1321–1331. https://doi.org/10.1016/j.visres.2003.10.024

Evans, K. K., & Treisman, A. (2005). Perception of objects in natural scenes: is it really attention free? *Journal of Experimental Psychology: Human Perception and Performance*, 31(6), 1476-1492.

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39, 175-191.

Felsen, G., & Dan, Y. (2005). A natural approach to studying vision. *Nature Neuroscience*, 8, 1643-1646. https://doi.org/10.1038/nn1608.

Ferrari, V., De Cesarei, A., Mastria, S., Lugli, L., Baroni, G., Nicoletti, R., & Codispoti, M. (2016). Novelty and emotion: Pupillary and cortical responses during viewing of natural scenes. *Biological Psychology*, 113, 75–82. https://doi.org/10.1016/j.biopsycho.2015.11.008

Ferrari V, Bradley MM, Codispoti M, Karlsson M, Lang PJ. (2013). Repetition and brain potentials when recognizing natural scenes: task and emotion differences. *Social Cognitive and Affective Neuroscience, 8*(8):847-54. https://doi.org/10.1093/scan/nss081

Ferrari, V., Mastria, S., & Codispoti, M. (2020). The interplay between attention and long-term memory in affective habituation. *Psychophysiology,* 020;57:e13572. https://doi.org/10.1111/psyp.13572

Ferrari, V., Bradley, M. M., Codispoti, M., & Lang, P. J. (2011). Repetitive exposure: Brain and reflex measures of emotion and attention. *Psychophysiology*, 48, 515–522. https://doi.org/10.1111/j.1469-8986.2010.01083.x

Ferrari, V., Codispoti, M., Cardinale, R., & Bradley, M. M. (2008). Directed and motivated attention during processing of natural scenes. *Journal of Cognitive Neuroscience*, 20, 1753–1761. https://doi.org/10.1162/jocn.2008.20121

Flaisch, T., Junghöfer, M., Bradley, M.M., Schupp, H.T., & Lang, P.J., (2008). Rapid picture processing: affective primes and targets. *Psychophysiology* 45, 1–10.

Folk, C. L. (2015). Controlling Spatial Attention: Lessons from the Lab and Implications for Everyday Life. In J. M. Fawcett, E. F. Risko, & A. Kingstone, *The Handbook of Attention* (pp.3-25). MIT press

Friston, K., Kiebel, S., 2009. Predictive coding under the free-energy principle. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 364 (1521), 1211–1221. https://doi.org/10.1098/rstb.2008.0300

Friston, K. (2010). The free-energy principle: a unified brain theory? *Nature Review Neuroscience*, 11, 127–138.

Galloway, D. W. (1946). An experimental investigation of structural lag in perception. *American Psychologist*, 1, 450 (Abstract).

Gibson, J. J. (1978). The ecological approach to visual perception. Houghton Mifflin.

Gollin, E.S. (1960). Developmental studies of visual recognition of incomplete objects. Perceptual and Motor Skills 11, 289–294. https://doi.org/10.2466/pms.1960.11.3.289.

Grassini, S., Holm, S. K., Railo, H., and Koivisto, M. (2016). Who is afraid of the invisible snake? Subjective visual awareness modulates posterior brain activity for evolutionarily threatening stimuli. *Biological Psychology*. 121, 53–61. doi: 10.1016/j.biopsycho.2016.10.007

Gregory, R. L. (1970). The intelligent eye. Weidenfeld & Nicolson.

Grossberg, S. (1980). How does a brain build a cognitive code? *Psychological Review,* 87, 1–51. https://doi.org/10.1037/0033-295X.87.1.1

Hajcak, G., Weinberg, A., MacNamara, A., & Foti, D. (2012). ERPs and the study of emotion. In: E.S. Kappenman, S.J. Luck (Eds.), *The Oxford handbook of event-related potential components*, 441-474.

Harris, C. R., & Pashler, H. (2004). Attention and the processing of emotional words and names: Not so special after all. *Psychological Science*, 15, 171–178. https://doi.org/10.1111/j.0956-7976.2004.01503005.x

Hedger, N., Adams, W. J., & Garner, M. (2015). Autonomic arousal and attentional orienting to visual threat are predicted by awareness. *Journal of Experimental Psychology. Human Perception and Performance*, 41(3), 798–806. http://doi.org/10.1037/xhp0000051

Hedger, N., H, L., Garner, M., & Adams, W. J. (2016). Are Visual Threats Prioritized Without Awareness? A Critical Review and Meta-Analysis Involving 3 Behavioral Paradigms and 2696 Observers. Psychological Bulletin, No Pagination Specified. http://doi.org/10.1037/bul0000054

Helmholtz H. von (1867), Handbook of Physiological Optics, vol. iii, Leopold Voss, Leipzig.

Joubert, O.R., Rousselet, G.A., Fabre-Thorpe, M., and Fize, D. (2009). Rapid visual categorization of natural scene contexts with equalized amplitude spectrum and increasing phase noise. *Journal of Vision* 9, 1-16. https://doi.org/10.1167/9.1.2

Junghofer, M., Elbert, T., Tucker, D. M., & Rockstroh, B. (2000). Statistical control of artifacts in dense array EEG/MEG studies. *Psychophysiology*, 37, 523–532. https://doi.org/10.1111/1469-8986.3740523

Keil, A., Bradley, M. M., Hauk, O., Rockstroh, B., Elbert, T., & Lang, P. J. (2002). Large-scale neural correlates of affective picture processing. *Psychophysiology*, 39, 641–649.

Keil, A., Moratti, S., Sabatinelli, D., Bradley, M.M., Lang, P.J., 2005. Additive effects of emotional content and spatial selective attention on electrocortical facilitation. *Cerebral Cortex*, 15, 1187–1197.

Lähteenmäki, M., Hyönä, J., Koivisto, M., & Nummenmaa, L. (2015). Affective processing requires awareness. *Journal of Experimental Psychology: General*, 144, 339–365. doi:10.1037/xge0000040

Lang, P., & Bradley, M. (2010). Emotion and the motivational brain. *Biological Psychology*, 84, 437–450.

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (1997). Motivated attention: Affect, activation, and action. In P. Lang, R. Simons, & M. Balaban (Eds.), Attention and orienting (p. 97-135). Mahwah, NJ: Erlbaum.

Lang, P. J., Greenwald, M. K., Bradley, M. M., & Hamm, A. O. (1993). Looking at pictures: Affective, facial, visceral, and behavioral reactions. *Psychophysiology*, 30(3), 261–273. https://doi.org/10.1111/j.1469-8986.1993.tb03352.x

Lang, P.J., Bradley, M.M., and Cuthbert, B.N. (2015). International affective picture system (IAPS): affective ratings of pictures and instruction manual. *Technical Report A-8.* Gainesville, FL: University of Florida.

Luo, C. R., & Snodgrass, G. (1994). Competitive activation model of perceptual interference in picture and word identification. *Journal of Experimental Psychology: Human Perception & Performance*, 20, 50-60.

MacNamara, A. (2018). In the mind's eye: The late positive potential to negative and neutral mental imagery and intolerance of uncertainty. *Psychophysiology*, 55(5), e13024. https://doi.org/10.1111/psyp.13024

Micucci, M. Ferrari, V., De Cesarei, A., & Codispoti, M. (2020). Contextual Modulation of Emotional Distraction: Attentional Capture and Motivational Significance. *Journal of Cognitive Neuroscience,* 32(4), 621-633. https://doi.org/10.1162/jocn_a_01505

Moors, A. (2007). Can cognitive methods be used to study the unique aspect of emotion: An appraisal theorist's answer. *Cognition & Emotion*, 21, 1238–1269. https://doi.org/10.1080/02699930701438061

Nickerson, R. S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology*, 2, 175—220.

Nummenmaa, L., Hyönä, J., & Calvo, M. G. (2010). Semantic categorization precedes affective evaluation of visual scenes. *Journal of Experimental Psychology: General,* 139, 222–246. https://doi.org/10.1037/a0018858

Öhman, A. (2005). The role of the amygdala in human fear: Automatic detection of threat. *Psychoneuroendocrinology*, 30, 953–958. https://doi.org/10.1016/j.psyneuen.2005.03.019

Peira, N., Golkar, A., Öhman, A., Anders, S., & Wiens, S. (2012). Emotional responses in spider fear are closely related to picture awareness. *Cognition and Emotion*, 26, 252–260.

Pessoa, L. (2005). To what extent are emotional visual stimuli processed without attention and awareness? *Current Opinion in Neurobiology*, 15, 188–196. https://doi.org/10.1016/j.conb.2005.03.002

Pessoa, L. (2008). On the relationship between emotion and cognition. *Nature Reviews. Neuroscience*, 9, 148–58. https://doi.org/10.1038/nrn2317

Peyk, P., De Cesarei, A., and Junghöfer, M. (2011), ElectroMagnetoEncephalography software: overview and integration with other EEG/MEG toolboxes. *Computational Intelligence and. Neuroscience* 2011: 861705. https://doi.org/10.1155/2011/861705

Rao, R.P. & Ballard, D.H. (1999). Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nature Neuroscience*. 2, 79–87

Sabatinelli, D., Keil, A., Frank, D.W.., & Lang, P.J. (2013). Emotional perception: Correspondence of early and late event-related potentials with cortical and subcortical functional MRI. *Biological Psychology*, 92(3), 513–519. http://dx.doi.org/10.1016/j.biopsycho.2012.04.005

Sanocki, T. (1993). Time course of object identification: Evidence for a global-to-local contingency. *Journal of Experimental Psychology: Human Perception and Performance*,19, 878-898.

Schendan, H. E., & Ganis, G. (2015). Top-down modulation of visual processing and knowledge after 250 ms supports object constancy of category decisions. *Frontiers in psychology*, 6, 1289. https://doi.org/10.3389/fpsyg.2015.01289

Schlögl, A., Keinrath, C., Zimmermann, D., Scherer, R., Leeb, R., and Pfurtscheller, G. (2007). A fully automated correction method of EOG artifacts in EEG recordings. *Clinical Neurophysiology* 118, 98–104. https://doi.org/10.1016/j.clinph.2006.09.003

Schupp, H. T., Schmälzle, R., Flaisch, T., Weike, A. I., & Hamm, A. O. (2012). Affective picture processing as a function of preceding picture valence: An ERP analysis. *Biological Psychology*, 91, 81-87. https://doi.org/10.1016/j.biopsycho.2013.02.002

Schupp, H. T., Stockburger, J., Schmälzle, R., Bublatzky, F., Weike, A. I., & Hamm, A. O. (2008). Visual noise effects on emotion perception: brain potentials and stimulus identification. *Neuroreport*, 19, 167-171. https://doi.org/10.1097/WNR.0b013e3282f4aa42

Schupp, H.T., Flaisch, T., Stockburger, J., and Junghöfer, M. (2006). Emotion and attention: event-related brain potential studies. *Progress in Brain Research* 156, 31–51. https://doi.org/10.1016/S0079-6123(06)56002-9

Schupp, H. T., Cuthbert, B. N., Bradley, M. M., Hillman, C. H., Hamm, A. O., & Lang, P. J. (2004). Brain processes in emotional perception: Motivated attention. *Cognition & Emotion*, 18, 593–611.

Snodgrass, J. G., & Hirshman, E. (1991). Theoretical explorations of the Bruner-Potter (1964) interference effect. *Journal of Memory and Language*, 30, 273-293.

Thorpe, S. J., Gegenfurtner, K. R., Fabre-Thorpe, M., & Bülthoff, H. H. (2001). Detection of animals in natural images using far peripheral vision. *European Journal of Neuroscience,* 14, 869–876.

Torralba, A., Oliva, A., 2003. Statistics of natural image categories. *Network: computation in neural systems*,14 (3), 391–412, http://dx.doi.org/10.1088/0954-898X 14 3 302.

VanRullen, R. (2011). Four common conceptual fallacies in mapping the time course of recognition. *Frontiers in Psychology*. 2:365. doi: 10.3389/fpsyg.2011.00365

Whitson, J. a, & Galinsky, A. D. (2008). Lacking control increases illusory pattern perception. *Science* (New York, N.Y.), 322, 115–7. https://doi.org/10.1126/science.1159845

Wichmann, F. A., Braun, D. I., & Gegenfurtner, K. R. (2006). Phase noise and the classification of natural images. *Vision research, 46,* 1520-1529.

Willenbockel, V., Sadr, J., Fiset, D., Horne, G.O., Gosselin, F., and Tanaka, J.W. (2010). Controlling low-level image properties: the SHINE toolbox. *Behavior Research and Methods 42*, 671-684. https://doi.org/10.3758/BRM.42.3.671

Wyatt, D. F, & Campbell, D. T. (1951). On the liability of stereotype or hypothesis. *Journal of Abnormal and Social Psychology*, 46, 496-500.

Footnotes

1. While in both experiments the affective modulation of the LPP in the 55% phase-scrambling condition was delayed compared to the more intact conditions, in Experiment 1 no emotional LPP modulation was observed until 800 ms, and only pleasant scenes differed from neutral scenes in the 800-1200 time interval; in Experiment 2, the unpleasant-neutral difference was delayed until 800 ms, while the LPP in response to pleasant scenes differed from that for neutral ones beginning from 400 ms after scene onset. However, when we performed an ANOVA with design Experiment x Time x Emotional Category, no significant interaction of Experiment with any of the other factors was observed, and in particular no Experiment x Time x Emotional Category interaction was observed, $F(2, 76) = .865$, $p = .424$, $\eta^2_p = .02$ in the 55% condition. Similarly, an ANOVA with design Experiment x Time x Degradation x Emotional Category was performed on the two phase scrambling conditions which are present in both experiments (55% and 0%), no Experiment x Time x Degradation x Emotional Category interaction was observed, $F(2, 76) = .720$, $p = .483$, $\eta^2_p =.019$.

Figure Captions

Figure 1 The top panel represents the sequence of picture presentation in the mixed and sequential blocks. The bottom left panel represents the sequence of events for each picture in the sequence.

Figure 2. The effects of phase scrambling on Categorization accuracy, separately for the mixed and the sequential condition. Error bars represent within-subject SEM. For descriptive purposes, a cumulative Weibull function is fitted to the data.

Figure 3. On the top row, waveforms indicate the effects of Picture and Phase scrambling on the LPP in Experiment 1, averaged over the central sensor group. The inset on the bottom left represents an overview of the sensors selected for scoring the LPP, plotted over a realistic 3D head model. On the bottom right, waveforms corresponding to the interaction of procedure and phase scrambling are reported.

Figure 4. Bar plots and scalp topographies (top view) show the LPP affective modulation in the time intervals 400-800 and 800-1200 ms, separately for each phase scrambling condition. For bar graphs, error bars represent within-subject SEM.

Figure 5. Effects of Picture and Phase scrambling on the LPP in Experiment 2, averaged over the central sensor group.

Figure 6. Bar plots and scalp topographies show the LPP affective modulation in the time intervals 400-800 and 800-1200 ms, separately for each phase scrambling condition. For bar graphs, error bars represent within-subject SEM.

Figure 7. The effects of phase scrambling on LPP affective modulation (emotional – neutral) in the 800-1200 time interval. Error bars represent within-subject SEM. For descriptive purposes, a cumulative Weibull function is fitted to the data.

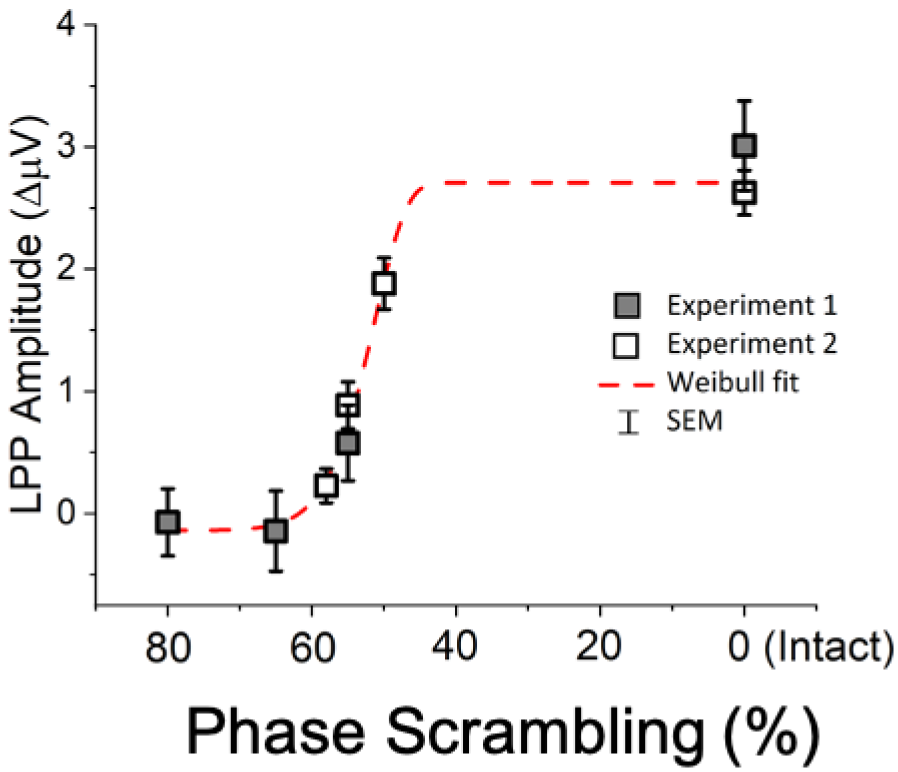
Tables

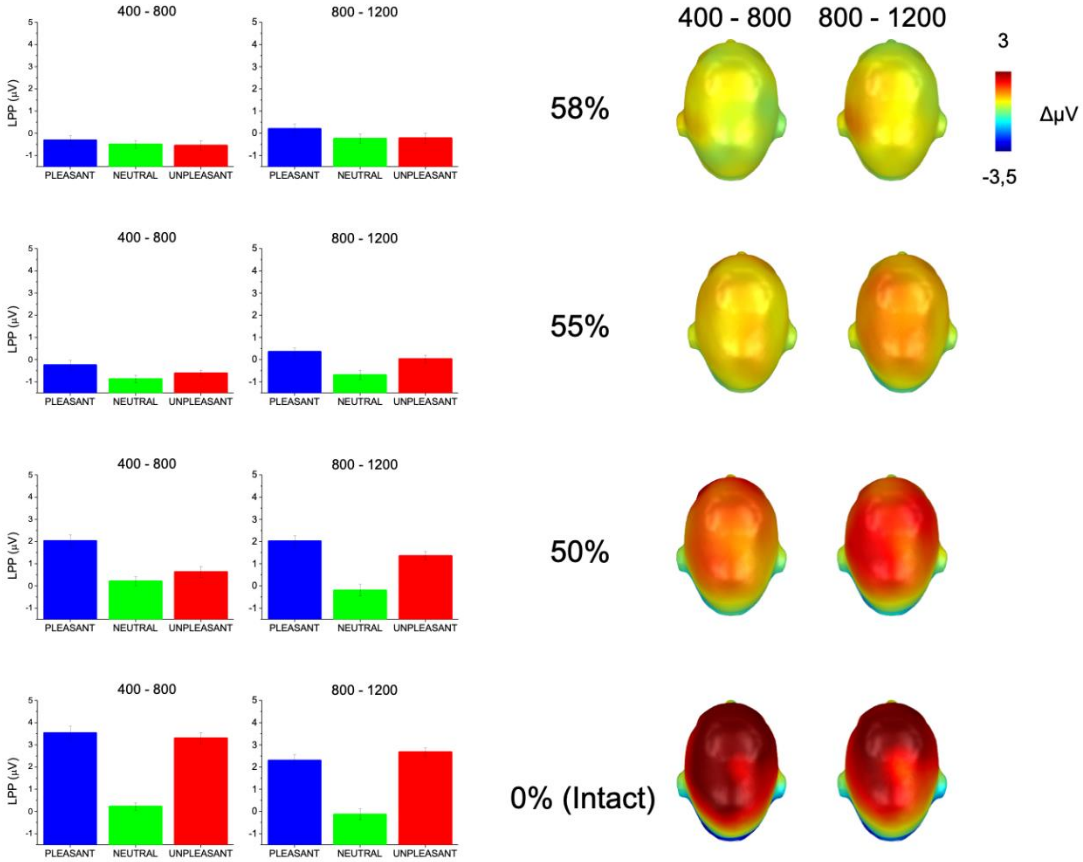
Table 1. Mean percentage of categorization accuracy and within-participants SEM (in parentheses) in Experiment 1, for each picture content, phase-scrambling condition, and procedure.

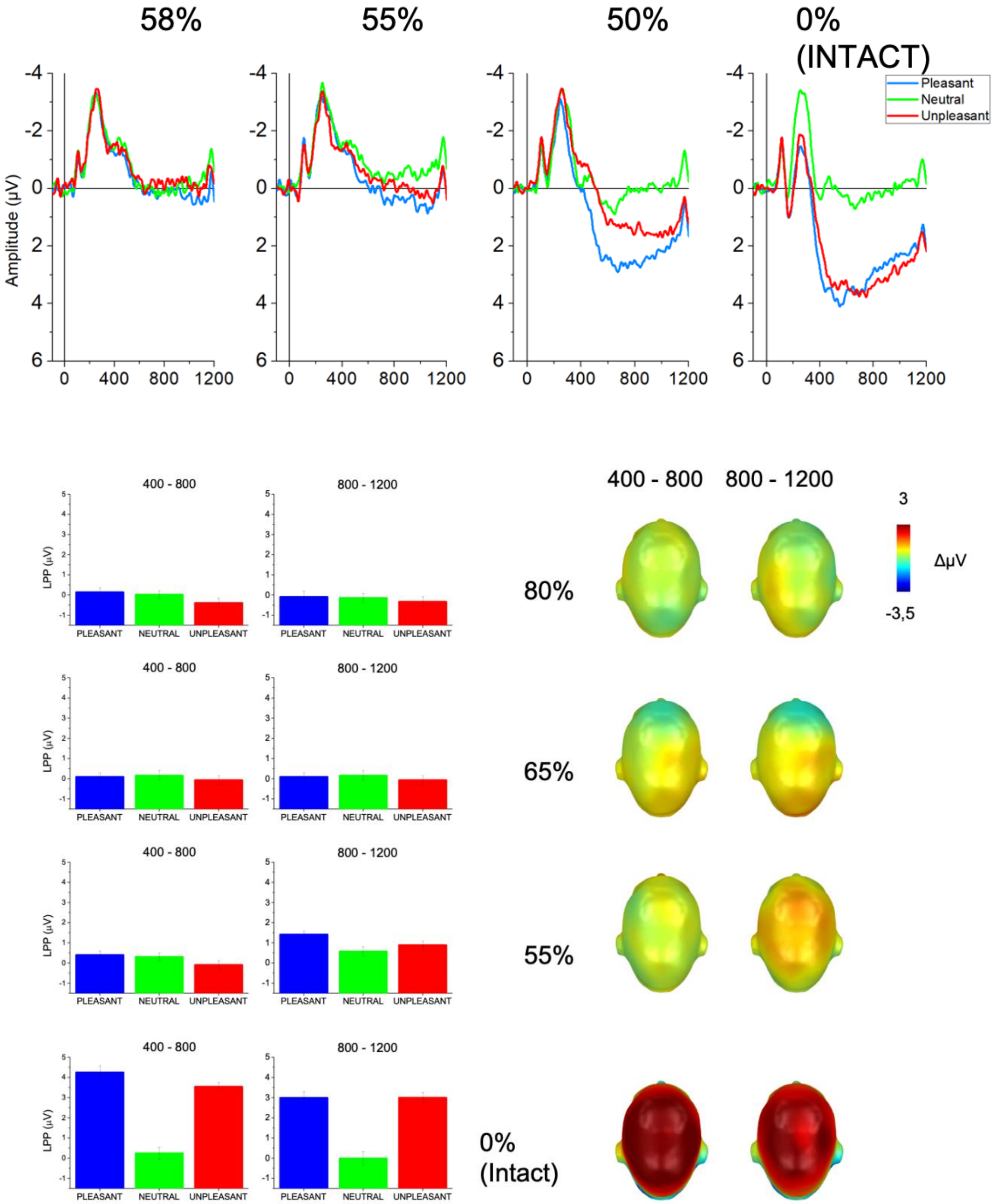
		PHASE SCRAMBLING				
	EMOTIONAL CATEGORY	80% (Most Degraded)	65%	55%	0% (Intact)	TOTAL
MIXED	Pleasant	48.15	59.5	88.75	100.00	74.1
		(2.48)	(3.95)	(3.04)	(<.01)	(5.47)
	Neutral	54.05	64.25	90.5	98.75	76.89
		(2.26)	(3.3)	(1.97)	(0.99)	(4.69)
	Unpleasant	47.6	49.75	70.5	98.6	66.61
		(3.45)	(3.25)	(4.11)	(0.83)	(5.57)
	SUBTOTAL MIXED	49.93	57.83	83.25	99.12	72.53
		(2.85)	(3.77)	(3.75)	(0.76)	(5.35)
SEQUENTIAL	Pleasant	44.5	57.75	86.75	99.5	72.13
		(2.93)	(3.25)	(2.77)	(0.34)	(5.56)
	Neutral	48.75	65.25	91.75	100.00	76.44
		(3.39)	(2.77)	(2.13)	(<.01)	(5.19)
	Unpleasant	49.25	48.75	62.25	99.00	64.81
		(3.89)	(3.16)	(3.68)	(0.57)	(5.54)
	SUBTOTAL SEQUENTIAL	47.5	57.25	80.25	99.5	71.13
		(3.46)	(3.42)	(4.11)	(0.39)	(5.54)
TOTAL		48.72	57.54	81.75	99.31	71.83
		(3.18)	(3.6)	(3.95)	(0.61)	(5.45)

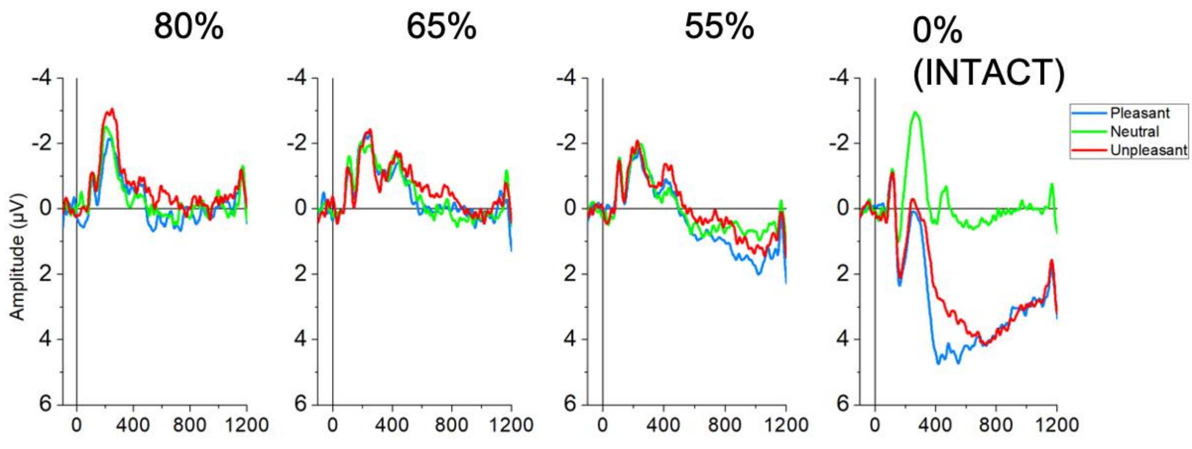
Table 2. Mean percentage of categorization accuracy and within-participants SEM (in parentheses) in Experiment 2, for each picture content, phase-scrambling condition, and procedure.

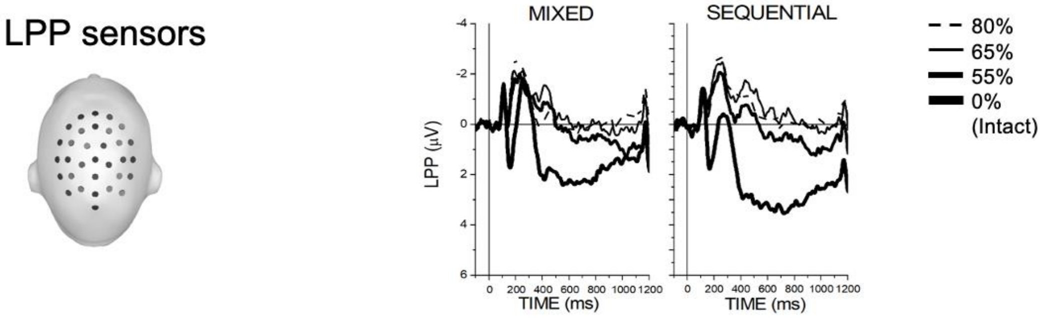
		PHASE SCRAMBLING				
	EMOTIONAL CATEGORY	80% (Most Degraded)	65%	55%	0% (Intact)	TOTAL
MIXED	Pleasant	77.75	91.25	99.25	99.75	92
		(3.25)	(1.41)	(0.53)	(0.24)	(2.68)
	Neutral	85.25	95.5	99.25	99.25	94.81
		(1.89)	(0.7)	(0.4)	(0.4)	(1.65)
	Unpleasant	63.25	79.25	93.5	99.5	83.88
		(2.48)	(2.92)	(1.46)	(0.49)	(3.75)
	SUBTOTAL MIXED	75.42	88.67	97.33	99.50	90.23
		(3.31)	(2.46)	(1.11)	(0.39)	(3.01)
SEQUENTIAL	Pleasant	81.00	88.25	98.25	99.75	91.81
		(2.14)	(1.78)	(0.53)	(0.24)	(2.22)
	Neutral	86.5	93.5	99.25	100.00	94.81
		(2.03)	(1.37)	(0.4)	(<.01)	(1.74)
	Unpleasant	61.5	68.5	88.25	99.00	79.31
		(3.52)	(2.48)	(1.74)	(0.45)	(4.09)
	SUBTOTAL SEQUENTIAL	76.33	83.42	95.25	99.58	88.65
		(3.58)	(3.09)	(1.55)	(0.31)	(3.24)
TOTAL		75.88	86.04	96.29	99.54	89.44
		(3.44)	(2.85)	(1.37)	(0.35)	(3.13)

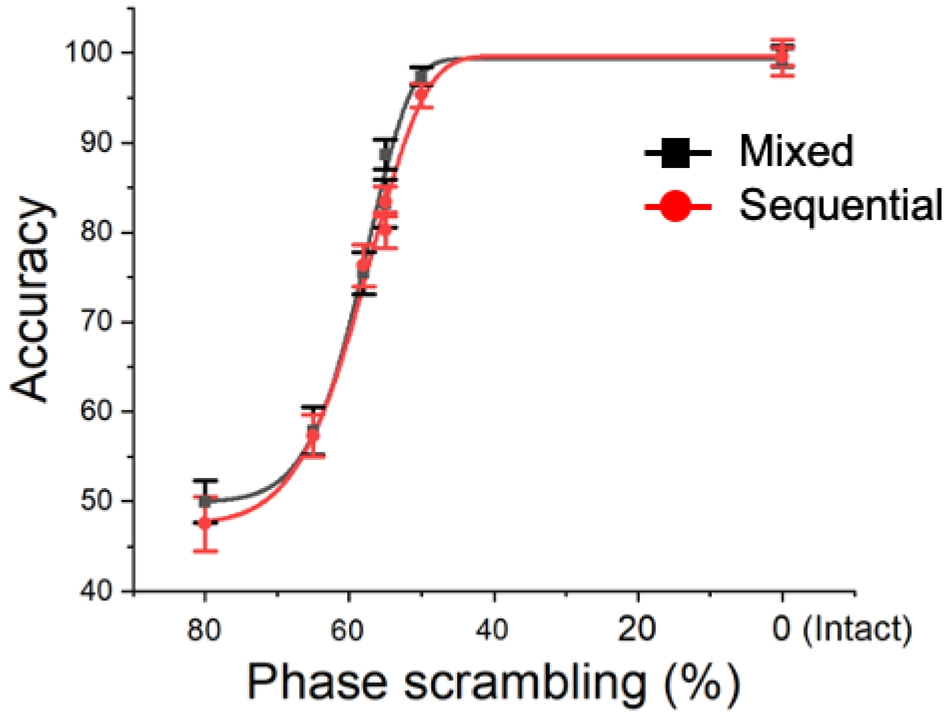

4
3
2
1
0
LPP Amplitude (ΔμV)
Experiment 1
Experiment 2
Weibull fit
SEM
80
60
40
20
0 (Intact)
Phase Scrambling (%)

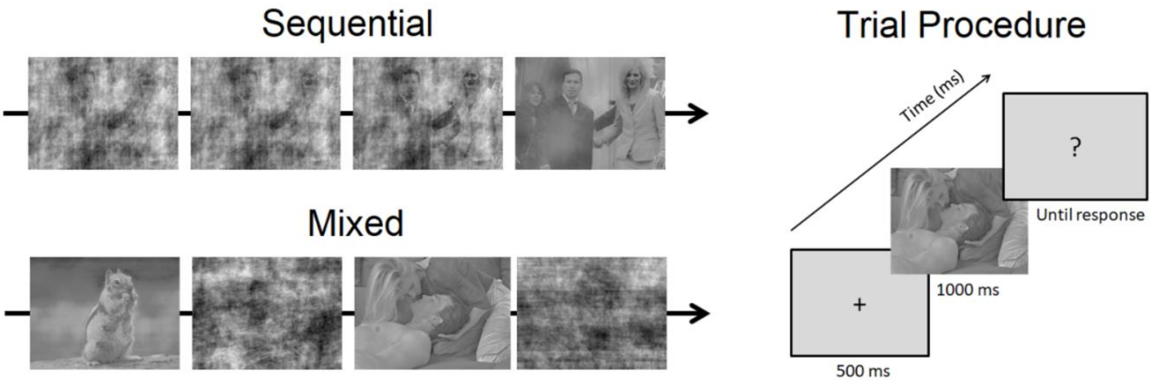

400 - 800
800 - 1200
LPP (μV)
PLEASANT
NEUTRAL
UNPLEASANT
58%
55%
50%
0% (Intact)
3
ΔμV
-3,5


58%
55%
50%
0% (INTACT)
Pleasant
Neutral
Unpleasant
Amplitude (µV)
400 - 800
800 - 1200
LPP (µV)
PLEASANT
NEUTRAL
UNPLEASANT
80%
65%
55%
0% (Intact)
3
ΔµV
-3,5


80%
65%
55%
0%
(INTACT)
Pleasant
Neutral
Unpleasant
Amplitude (µV)
-4
-2
0
2
4
6
0
400
800
1200


LPP sensors
MIXED
SEQUENTIAL
80%
65%
55%
0%
(Intact)
LPP (µV)
TIME (ms)
0
200
400
600
800
1000
1200


Mixed
Sequential
Accuracy
40
50
60
70
80
90
100
80
60
40
20
0 (Intact)
Phase scrambling (%)


Sequential
Mixed
Trial Procedure
Time (ms)
+
500 ms
1000 ms
?
Until response