Article:

Bekhtereva, V, Craddock, M and Müller, MM (2015) Attentional bias to affective faces and complex IAPS images in early visual cortex follows emotional cue extraction. NeuroImage, 112. 254 - 266. ISSN 1053-8119

https://doi.org/10.1016/j.neuroimage.2015.03.052

Attentional bias to affective faces and complex IAPS images in early visual cortex follows emotional cue extraction

Valeria Bekhtereva[1], Matt Craddock[2], Matthias M. Müller[1]

[1] Institute of Psychology, University of Leipzig, Leipzig, Germany

[2] Institute of Psychological Sciences, University of Leeds, Leeds, UK

Corresponding author:

Matthias M. Müller

Institute of Psychology

University of Leipzig

Neumarkt 9-19

04109, Leipzig, Germany

Tel. ++49 (0)341-973 9543,

Fax: ++49 (0)341-973 9269,

e-mail: m.mueller@rz.uni-leipzig.de

Abstract

Emotionally arousing stimuli are known to rapidly draw the brain's processing resources, even when they are task-irrelevant. The steady-state visual evoked potential (SSVEP) response, a neural response to a flickering stimulus which effectively allows measurement of the processing resources devoted to that stimulus, has been used to examine this process of attentional shifting. Previous studies have used a task in which participants detected periods of coherent motion in flickering random dot kinematograms (RDKs) which generate an SSVEP, and found that task-irrelevant emotional stimuli rapidly withdraw attentional resources from the task-relevant RDKs. However, it is not clear whether the changes in the SSVEP response are conditional on higher-level extraction of emotional cues as indexed by well-known event-related potential (ERPs) components, or if affective bias in competition for visual attention resources could be a consequence of an inherent, relatively time-invariant shifting process. In the present study, we used two different types of emotional distractors - IAPS pictures and facial expressions – for which emotional cue extraction occurs at different speeds, being typically earlier for faces (at ~170 ms, as indexed by the N170) than for IAPS images (220-230 ms, Early Posterior Negativity, EPN). We found that attentional resources were withdrawn from the foreground task towards task-irrelevant emotional background images from the International Affective Picture System (IAPS) following the extraction of emotional cues as indexed by visual ERP components. We also found that emotional modulation of attentional resources as measured by the SSVEP occurred earlier for faces (around 180 ms) than for IAPS pictures (around 400 ms). This is consistent with low-level attentional resources being re-allocated after emotional cue extraction rather than being linked to a time-fixed shifting process.

Keywords: Human EEG, Steady-state visual evoked potentials (SSVEPs), attention-emotion interaction, temporal dynamics of competition for processing resources, N170, EPN

1. Introduction

The extraction of relevant visual information is pivotal for adaptive behavior. It is generally accepted that, due to limited processing capacity, stimuli need to compete for sensory processing resources (Desimone and Duncan 1995; Desimone 1998; Kastner et al. 1998; Kastner and Ungerleider 2001; Petersen and Posner 2012). In particular, emotional stimuli can bias visual processing resources and attract attention at the expense of other, emotionally neutral stimuli (Bradley et al. 2003; Keil et al. 2005; Lang and Davis 2006). For example, the threat posed by a knife put to somebody's throat or by a gun directed to the observer rapidly captures attention, presumably due to the intrinsic biological relevance of the information such a scene conveys. Although a number of studies support this notion (Anderson et al. 2003; Öhman et al. 2001; Vuilleumier et al. 2001), others nevertheless suggest that at least some additional attention needs to be allocated to emotional stimuli in order to fully capture attentional processing resources (cf. Holmes et al. 2006; Pessoa et al. 2002). For example, Pessoa (Pessoa et al. 2002) found that emotional compared to neutral faces are not linked to increased amygdala activation when subjects attended to a different location than the one of face presentation. However, as of today, it appears not to be a question of "all or nothing": the picture becomes more fine-tuned with the acknowledgment of a number of factors such as task-relevance and type of the emotional stimulus presented (i.e. emotional faces or complex emotional images), its valence (pleasant or unpleasant) and the elicited arousal of an emotional stimulus (for a review see Okon-Singer et al. 2013).

In recent years, we published a number of studies in which we used a distraction paradigm to investigate the neural dynamics of attentional resource competition in early visual cortex between a foreground task and an emotional distractor (Hindi Attar et al. 2010; Müller et al. 2008; Müller et al. 2011; Schönwald and Müller 2014). In these experiments, participants attended to a random dot kinematogram (RDK) that continuously flickered at a specific frequency, and were asked to detect and respond to periods of coherent motion events of those dots (targets). At some point during the trial, a task-irrelevant neutral or emotional image from the International Affective Picture Set (IAPS, Lang et al. 2008) appeared in the background. Subjects were instructed to ignore these background images. The flickering RDKs that formed the foreground task elicited the steady-state visual evoked potential (SSVEP), a continuous oscillatory response with the same frequency as the driving stimulus (Regan 1989). Importantly, SSVEP amplitude increases significantly when a flickering stimulus is attended compared to when that stimulus needs to be ignored (Morgan 1996; Müller et al. 1998; Müller et al. 2006). The generators of the SSVEP to low-level stimuli that are presented at higher frequency rates have consistently been found in early visual areas, including primary visual cortex (Di Russo et al. 2007; Müller et al. 2006). Thus, the SSVEP of low- level stimuli flickering at high frequency rates allows investigation of neural attention-emotion interactions at an early sensory processing stage. Given its ongoing oscillatory nature, the SSVEP serves as a powerful objective electrophysiological measure to investigate temporal neural dynamics of competitive interactions between attention and emotional stimuli.

In our previous studies we consistently showed a significantly greater decrease in SSVEP amplitude for an emotional compared to a neutral background IAPS picture between 400 to 475 ms following the picture onset (Hindi Attar et al. 2010; Müller et al. 2011; Schönwald and Müller 2014) and lasting up to 1 sec (Müller et al. 2008). In the distraction paradigm, the reduction of SSVEP amplitude as a function of valence of the background distractor signifies the withdrawal of attentional resources away from the flickering stimulus toward a background distractor-image (i.e. foreground task; cf. Hindi Attar et al. 2010; Müller et al. 2008). Interestingly, this differential effect even occurred when we presented unpleasant or neutral IAPS for 200 ms only (Müller et al. 2011), which allowed emotional content categorization but no further elaborative processing (Larson et al. 2005; Schupp, Junghöfer et al. 2004). Thus, the attentional effects as measured by the SSVEP are relatively late when considering the rapidity associated with attentional capture by emotional stimuli.

One possibility is that competition for attentional resources in early visual cortex is preceded by emotional content extraction, and the rather slow biasing observed using the SSVEP might be due to feedback (Pourtois et al. 2013) or re-entrant mechanisms (Keil et al. 2009). The visual evoked potential (VEP) elicited by IAPS images offers a method to test this hypothesis. As a neural signature of emotional cue extraction, the early posterior negativity (EPN) was examined (Junghöfer et al. 2001; Schupp et al. 2003). The EPN is considered to be an early neural marker of affective stimulus discrimination, occurring 200-300 ms after stimulus onset as a more negative amplitude deflection for emotional than neutral images (Olofsson et al. 2008; Schupp et al. 2008; Wiens and Syrjänen 2013). It is not tied to a clear single peak but rather extends over a wider time range (Rellecke et al. 2012). It may index a “call for attentional resources” (cf. Wiens and Syrjänen 2013); for review: Schupp et al. 2006) and is affected by picture complexity (i.e. natural scenes vs. figures) and arousal level (Junghöfer et al. 2001; Schupp et al. 2003). Additionally, in some studies (see a recent overview by Olofsson et al. 2008), the effects of affective processing are seen even earlier at the level of the N1 (with a variation of the peak latency between 160 to 190 ms, Weinberg and Hajcak 2010). The topographical distribution of differences for emotional minus neutral pictures in the N1 and EPN is almost identical, and thus the neural signature of emotional cue extraction of IAPS images may be better described as early negativities (Weinberg and Hajcak 2010), or an N1-EPN complex (Schönwald and Müller 2014).

In analyzing both measures, i.e. the VEP elicited by IAPS images and the SSVEP evoked by the flickering squares or dots that formed the foreground task, we found that the IAPS N1-EPN complex (190-360 ms) clearly preceded the differential effect of SSVEP amplitude reduction as a function of valence of the background IAPS, which was observable around 400 ms (Schönwald and Müller 2014). Nevertheless, although the result strongly supported the need for cue extraction to occur before attentional biasing, as indexed by the SSVEP, it cannot entirely exclude an alternative explanation: the SSVEP effect may reflect an “inherent” shifting time that is needed to shift attentional resources to the background image after it appears. Consequently, the emotional SSVEP effect may occur after that fixed shifting interval, rather than as a direct consequence of emotional cue extraction. While a spatial shift seems very unlikely, given that RDKs and IAPS images were presented superimposed, such a time-invariant shifting interval could consist of two other processes.

The first such process might reflect an involuntary shift of attention driven by the onset of a new object, which produces a strong capture of attention (Yantis and Jonides 1984, 1996). In a previous study using the same paradigm as that used here, we found that the amplitude of the SSVEP significantly decreased around 130 ms after a phase-scrambled, meaningless image changed to a meaningful, concrete image depicting an emotional or neutral scene. In contrast, when phase-scrambled, meaningless images changed to other phase-scrambled meaningless images, there were no significant decreases in amplitude (Hindi Attar et al. 2010). Interestingly, the initial decrease in SSVEP amplitude does not differ across valence categories, unlike later changes in SSVEP amplitude. This “first deflection” likely reflects the consequence of fast visual categorization (Thorpe et al. 1996) or first “initial object classification” activity (cf. Schendan and Lucia 2010) on low-level processing (i.e. through feedback), thus providing evidence for discrimination between phase-scrambled vs. intact objects (Grill-Spector and Malach 2004). The latency of this effect parallels reports of enhanced amplitudes in early category-specific VEP components from 120 to 200 ms in response to intact familiar objects (i.e faces, cars) relative to phase-scrambled versions of these images (Rossion and Caharel 2011; Rousselet et al. 2008). Therefore, the presentation of a concrete IAPS image that contains at least one object can elicit an involuntary shift of attention (Schreij et al. 2008; Schreij et al. 2010).

A second shifting process might be related to certain features of the background image that may or may not be the relevant features that constitute “emotionality”. In a recent frequency tagging study using similar stimuli and methods, we found that SSVEP amplitudes increased significantly with non-spatial feature-based shifts at about 220 ms after the presentation of the shifting cue (Andersen and Müller 2010). Aggregating the approximate timing of these two processes – an involuntary capture by a new object and a non-spatial, feature-based shift – together, and assuming that they largely follow each other most certainly with some temporal overlap, suggests an approximate total shifting time of 350 ms. This coincides with the time point at which the differential SSVEP effect starts to evolve.

A possible way to test for the alternative explanation of a time-invariant shifting procedure is to present background distractors that are known for a more rapid emotional cue extraction, compared to complex IAPS images. A wealth of studies exist, too many to cite here, that use facial expressions to investigate attention-emotion interactions. Importantly in the current context, a number of EEG studies have demonstrated that emotional cue extraction for faces is much faster compared to IAPS images, with a latency of about 120-170 ms, i.e. with the face-specific N170 (Ashley et al. 2004; Batty and Taylor 2003; Blau et al. 2007; Eimer and Holmes 2007; Krombholz et al. 2007). The N170 component serves as an early neural marker signifying face perception (Bentin et al. 1996; Bentin and Deouell 2000) and has been reported to be more negative with emotional compared to neutral faces (Ashley et al. 2004; Batty and Taylor 2003; Blau et al. 2007). If the SSVEP is linked to a relatively time-invariant shifting process as outlined above, one would not expect a significant latency difference of the amplitude divergence for facial compared to IAPS distractors. If, however, the SSVEP effect follows emotional cue extraction, one would expect a shorter latency of SSVEP amplitude differences between emotional and neutral stimuli for faces than for IAPS background images, given that emotional cue extraction obviously occurs faster for faces.

Similar to our previous studies, the current study presented unpleasant or neutral IAPS scenes and fearful or neutral facial expressions as a background to a flickering RDK. Participants attended to the RDK, and were asked to detect and respond to periods of coherent motion. Since IAPS images and images of faces may differ on a variety of low-level properties such as complexity, luminance contrast, and spatial frequency content, we minimized such differences by equating the means and standard deviations of the luminance distributions as well as rotational average Fourier amplitudes for each spatial frequency across images between the two stimulus sets (Willenbockel et al. 2010). Our results clearly indicate that the withdrawal of attentional resources from the foreground task chronologically follows the emotional cue extraction of the background image and, thus, are supportive for the "following emotional cue extraction hypothesis" rather than being linked to relatively time-invariant shifting process.

2. Materials and Methods of Experiment 1: competition between attended task and distracter-images

2.1 Subjects

18 subjects (6 male) with a mean age of 24 years (standard deviation [SD] =4.68) with normal or corrected to normal vision participated in the experiment. Prior to recordings subjects received information about the study goals and gave written informed consent. All subjects received either class credits or monetary compensation for their participation (6 € per hour). The study conformed to the

Code of Ethics of the World Medical Association and the requirements for EEG studies of the local ethics committee of the University of Leipzig.

2.2. Stimuli

Task-stimuli constituted an array of 115 randomly distributed moving red squares (each 0.6° x 0.6° of visual angle) superimposed upon a grayscale fearful or neutral facial expression or an unpleasant or neutral IAPS picture. For a baseline measure, these images were first presented in a scrambled version (see Figure 1A). Size of images extended 13.3° x 10.71° of visual angle. A centrally located red cross served as fixation during the entire experiment. Scrambling of images was performed by a Fourier transform, randomizing phase-components of each image but preserving its amplitude spectrum. The outcome images had the same low-level physical characteristics (luminance, spectral energy) as their non-scrambled originals, while any content-related information was distorted.

60 fearful and 60 neutral facial expressions were taken from the FACES Life span Database of Facial Expressions (Ebner et al. 2010), with all posers of Caucasian origin (all items are available upon request). 60 unpleasant and 60 neutral picture scenes (any picture content with facial expressions was intentionally excluded) were selected from IAPS image set (Lang et al. 2008). Thus, the selected images comprised two valence categories for facial expressions and two valence categories for IAPS pictures, resulting in 4 experimental conditions. All stimuli were converted to a grayscale format with the MATLAB image processing toolbox. The mean (representative of the global luminance) and standard deviation (representative of RMS contrast) of the luminance distribution of every grayscale image was calculated on the intensity of pixels ranging from 0 to 1 grayscale value (with minimum 0 – black and maximum 1 - white). Furthermore, using SHINE toolbox (Willenbockel et al. 2010), all images were matched in their spatial frequencies (thus, redistributing luminance energy, so that each image had approximately the same luminance energy at each spatial frequency) between all of the experimental conditions to have equal luminance (M=0.44; SD=0.0004; $F_{3,296}$=0.127, p=0.94, partial η^2=0.001) and contrast (M=0.096; SD=0.0002; $F_{3,296}$=0.353, p=0.79, partial η^2=0.004) composition. In order to match the aspect ratio of the face image to that of IAPS images, we added gray borders based on the mean luminance of the background of the face image to its left and right. Additionally, we measured the luminance of red squares that on average was 30.17 cd/m², while the mean luminance of the background images was 28.40 cd/m².

2.3 Experimental procedure and design

Stimuli were presented against a black background on a 19 inch computer screen which was set at a resolution of 800x600 pixels and 16 bits per pixel color mode, with a refresh rate of 60 Hz and a viewing distance of 80 cm. Red squares were randomly distributed across centrally presented facial expressions/IAPS pictures at a flicker frequency of 15 Hz. As shown previously, a tagging frequency of 15 Hz of the squares allows for the simultaneous extraction of the VEP elicited by the presentation of the background face or IAPS picture and the analysis of the time course of SSVEP amplitudes(Schönwald and Müller 2014). Squares were presented 'on' for 2 frames and 'off' for 2 frames to elicit the steady state visual evoked potential (SSVEP). One trial lasted for 4533 ms, resulting in 68 cycles of 15 Hz stimulation. After each trial a black screen with the fixation cross was presented for an interval that varied randomly between 1150 and 1650 ms. The timing and order of image presentation was controlled using the Cogent toolbox for Matlab (Cogent, www.vislab.ucl.ac.uk/Cogent/; The Mathworks, Inc, Natick, Massachusetts).

Each trial began with the simultaneous onset of a scrambled version of a background IAPS picture/facial expression superimposed with flickering red squares. At a certain time point during a trial that scrambled facial expression/IAPS picture changed to normal intact view of the same image. Every 16.67 ms (i.e. 1 frame of 60 Hz refresh rate of the screen) the dots moved randomly: either up, down, left or right by 0.04 degrees of visual angle to produce a Brownian motion. Targets were defined as events in which 35 % of the dots moved coherently in one of the four cardinal directions for 4 frames of the stimulation frequency of 15 Hz (i.e. for 266.68ms). During one trial between 0 and 4 coherent motion targets were possible. Critically, in order to analyze the time course of behavioral data, target events were uniformly distributed across each single cycle of the 15 Hz flicker stimulation (66.7 ms), such that over the entire experiment, four target events occurred in each time window in every experimental condition.

Participants were instructed to detect these events as quickly and accurately as possible by pressing the space bar on the keyboard while ignoring the background images as task-irrelevant. Halfway through the experiment the responding hand was changed with the starting hand counterbalanced across subjects. We presented 600 trials that were subdivided into 10 blocks with 60 trials each. In order to exclude any anticipation effects for the time of face/image change, the change occurred randomly in an early (200 – 733 ms, 7 % of trials), middle (1267 – 2400 ms, 80 % of trials), or late (2467 – 4067 ms, 13 % of trials) time window after trial onset. Trials with early and late face/image changes served as 'catch trials' and were excluded from further analysis (120 trials in total). For these trials a different set of facial expressions and IAPS pictures was chosen from the ones we used in trials that were included in the data analysis. Thus, there were 480 trials included in the final data analysis (120 per experimental condition). Each image or face was presented twice throughout the experiment (the repetition of the same image never occurred within one block) and randomized such that no images from the same condition occurred within three consecutive trials. After every 60 trials, i.e. one block, subjects were allowed to take a break. During this time, participants received feedback of their behavioral performance – the mean percentage target hit-rate and the mean reaction times for the block. Subjects pressed the space bar to continue with the next block. All subjects performed a short training session (up to 3 blocks) on the recording day until they reached a stable performance of about 60%, after which the actual experimental session started. We used a different set of images for this practice session.

After the EEG session, participants were asked to view the images of facial expressions and IAPS pictures used in the experiment in randomized order and to rate them on the dimension of affective arousal and valence on the 9-point Self-Assessment-Manikin (SAM) scale (Bradley and Lang 1994).

2.4 Electroencephalography recordings

Brain electrical activity was recorded from 64 Ag/AgCl scalp electrodes mounted in an elastic cap using a BioSemi ActiveTwo system (BioSemi, The Netherlands) with a sampling rate of 256 Hz. Lateral eye movements were monitored with a horizontal electroocculogram (EOG), while vertical eye movements and blinks were monitored with a bipolar montage positioned below and above the right eye (vertical EOG). Two additional electrodes were used as reference and ground (CMS – “Common Mode Sense” and DRL – “Driven Right Leg”; see http://www.biosemi.com/faq/cms&drl.htm).

----

insert Figure 1 about here

2.5 Data analysis: Behavioral data and SAM ratings

Correct responses between 250 and 1000 ms after target onset were considered as hits. Later button presses or no response were seen as misses. Onsets of target events were uniformly distributed over the time window from 800 ms before and 1867 ms after the onset of an intact background face/IAPS picture. Given that the duration of the event was 266.7 ms, we calculated correct responses within time bins of 266.7 ms by averaging across 4 short bins of 66.67ms (in order to keep every time bin identical to the duration of the target event (see also (Hindi Attar et al. 2010; Schönwald and Müller 2014)). This resulted in 10 time windows of 266.7 ms each. We focused further calculations on the time interval from 800 ms before up to 1334 ms after the face/picture change (3 time bins before and 5 after) based on the results of our previous studies, in which the same distraction paradigm was used (see above). All calculations were done in relation to the bin in which the onset of the target event occurred according to the uniform distribution of events over time. Repeated-measures ANOVAs with the within-subjects factors of Stimulus (IAPS, Face), Emotion (negative, neutral) and Time (bin 1-8) were calculated. To further test differences between conditions with respect to time bins, we conducted post-hoc t-tests using the Bonferroni correction for multiple comparisons where necessary.

Mean arousal and valence SAM ratings for fearful and neutral facial expressions, unpleasant and neutral IAPS pictures as well as comparisons between faces and IAPS were subjected to paired t-tests. Generalized eta (η_g) or Cohen's d (d) are reported as a measure of effect size (Baguley 2012a; Bakeman 2005; Olejnik and Algina 2003). All statistical analyses were conducted in R (R Core Team 2012).

2.6 Data Analysis: EEG

2.6.1 SSVEP analysis

EEG epochs between 1000 ms before and 2000 ms after picture change were extracted. Epochs were rejected manually when contaminated with muscle artifacts, blinks or eye movements. Automatic artifact detection was also performed using 'Statistical Control of Artifacts in Dense Array EEG/MEG studies' (Junghöfer et al. 2000). This procedure corrects artifacts such as noisy electrodes using a combination of channel approximation and epoch exclusion based on statistical parameters of the data. Epochs with more than 10 contaminated electrodes within a particular area of the electrode montage were excluded from further analysis. This resulted in an average rejection rate of 14.8% with no differences between conditions. Data was re-referenced to the average reference, and linear trends were removed. As a final step, all epochs were averaged for each individual and condition, respectively. Then the SSVEP amplitudes were extracted from the averaged epochs by means of a Gabor filter centered at 15 Hz with a frequency resolution of +-2.0 Hz full-width at half-maximum (FWHM) resulting in a temporal resolution of +/-110.3 ms. To identify electrodes with maximum SSVEP amplitudes for statistical analysis, we calculated the iso-contour voltage map averaged across all subjects and all conditions (see Figure 1B). The topographical distribution of SSVEP amplitude values was maximal at electrode Oz. This electrode was taken for each participant for further statistical analysis. Given the directionality of the hypothesis regarding SSVEP amplitudes based on the series of our previous studies with a similar experimental design (see Introduction) with a reduction of SSVEP amplitudes for emotional relative to neutral conditions, we used running one-tailed paired t-tests for the statistical analysis. Only time intervals with a minimum of 10 consecutive significant sampling points are reported (Andersen and Müller 2010). In addition to the statistical analysis of amplitude differences between emotional and neutral stimuli we conducted a direct test between face and IAPS background images. In a first step we tested for latency differences of the first SSVEP amplitude reduction after the presentation of the respective stimulus between face and IAPS stimuli. In a second step, we calculated difference curves (emotional minus neutral) for face and IAPS, respectively, and tested them against each other with running t-tests (one–sided t-tests were used due to the clear hypothesis of amplitude directions). Extraction of epochs from the continuous recording, re-referencing and scalp voltage maps were made in EEGLAB (Delorme and Makeig 2004) combined with in-house written scripts and running under the MATLAB (The Mathworks, Natick, MA).

2.6.2 Event-related potential analyses

We analyzed the P1 and N170 that were elicited by the faces as well as the N1, EPN and LPP components elicited by the IAPS images. To this end, continuous EEG data was bandpass filtered (Kaiser windowed FIR, 0.5-14 Hz; (Widmann 2006). We then extracted epochs from 133 ms before to 1300 ms after picture onset of the continuous EEG. The mean of the 133 ms before face/picture onset (i.e. 2 full cycles of 15 Hz stimulation) served as baseline and the mean was subtracted from each data point of the ERP epochs. Prior to artifact rejection, trials with horizontal eye movements >25 µV (about 1.5 degrees of visual angle) were removed. Overall 8.06 % of trials (437 out of 480 trials for 18 participants) were rejected from the analysis after artifact detection (Junghöfer et al. 2000) with no differences between conditions.

In order to identify peak latencies and relevant electrodes for statistical analysis of the respective ERP components, we averaged across emotional and neutral conditions within each stimulus category. Following that we calculated grand mean topographies of the respective components, based on visual inspection and previous results (Schönwald and Müller 2014), to identify electrodes with greatest amplitudes (see below and Figure 2). Data filtering, extraction of epochs, re-referencing to average reference and further analyses of ERP components were made in EEGLAB (Delorme and Makeig 2004) in combination with in-house scripts written in MATLAB (The Mathworks, Natick, MA) as well as an open-source Matlab toolbox (ERPLAB Toolbox, http://erpinfo.org/erplab). The following time windows to calculate mean amplitudes and electrodes were chosen for statistical analysis.

2.6.2.1 Faces

P1 (110 to 130 ms, +-10 ms around its peak latency at 120 ms), calculated at occipital electrodes O1/O2 where the positive deflection was most pronounced.

N170 (156 to 196 ms, ±20 ms around its peak latency at 176 ms), quantified at the parietal electrode sites P9, P10 where the maximal N170 deflection was observed (see Figure 2A).

2.6.2.2 IAPS

P1 (110 to 190 ms) quantified at O1 and O2. As can be seen from Figure 2B-C for N1-EPN complex, at around 110 and 190 ms the P1 component exhibited two peaks, most certainly caused by the 15 Hz SSVEP. Given this somewhat broader distribution, we decided to average across the broader time window for P1 analysis.

N1 (202-242 ms, ±20 ms around its peak latency at 222 ms; see Figure 2B), at parieto-occipital sites I1, I2, P7, P8, P9, P10, P07, PO8, PO3, PO4, O1 and O2 based on the selection of electrodes in the similarly designed previous study by Schönwald and Müller (Schönwald and Müller 2014).

For EPN analysis a time window from 276 to 426 ms was taken at the same cluster of 12 electrodes as for N1 (see Figure 2C and (Schönwald and Müller 2014)).

For all components we calculated the mean amplitude across the respective electrodes. These values were subjected to paired t-tests. Cohen's d (d) is reported as a measure of effect size.

3. Results

3.1 SAM ratings

As expected, fearful faces were rated as less pleasant (valence mean rating 3.92, SD=0.67, t_{17}=-6.0, p<0.001, d=-1.41) and more arousing (arousal mean rating 4.5, SD=1.80, t_{17}=5.4793, p<0.001, d=1.29) than neutral faces (valence mean rating 4.96, SD=0.27; arousal mean rating 2.62, SD=1.34). Similarly, unpleasant IAPS pictures were rated as less pleasant (valence mean rating 2.52, SD=0.45, t_{17}=-22.6, p<0.001, d=-5.32) and more arousing (arousal mean rating 6.14, SD=1.18, t_{17}=13.8, p<0.001, d=3.25) than neutral pictures (valence mean rating 5.6, SD=0.39; arousal mean rating 3.06, SD=1.11). Furthermore, faces were rated as less arousing than IAPS pictures: fearful faces were seen as less arousing than unpleasant IAPS (t_{17}=-5.35, p<0.001, d=-1.26), and neutral faces were less arousing than neutral IAPS (t_{17}=-2.52, p<0.05, d=-0.59). Also, valence was lower for unpleasant pictures than fearful faces (t_{17}=10.76, p<0.001, d=2.54), and higher for neutral faces than neutral IAPS (t_{17}=-5.90, p<0.001, d=-1.39).

3.2 ERP results

Figure 2 depicts the grand mean ERPs elicited by faces and IAPS images and the respective topographical distributions. The EPN is usually shown as a difference curve (emotional minus neutral). In order to allow for a better comparison to existing studies, we decided to present the topographical distribution of the N1 and EPN from the difference values. As noted in the introduction, the almost identical topographical distribution of the two components gives rise to the notion of an N1-EPN complex.

We found no significant differences in the P1 between fearful and neutral faces ($t_{17} = -0.82$, $p=0.42$, $d=-0.19$). Similar to previous studies cited in the Introduction section, the N170 amplitude was significantly more negative for fearful than neutral facial expressions ($t_{17} = -5.37$, $p<0.001$, $d=-1.26$). Similar to what we found for faces, there was no significant differences in the P1 between unpleasant and neutral pictures ($t_{17} = -1.13$, $p=0.28$, $d=-0.21$). However, the N1 significantly differentiated between unpleasant and neutral IAPS images ($t_{17} = -8.81$, $p<0.001$, $d=-2.07$) being less positive for unpleasant than for neutral IAPS pictures. That difference continued to the EPN time range, with lower amplitudes for emotional than neutral IAPS ($t_{17} = -4.12$, $p<0.001$, $d=-0.97$). Thus, the first difference between emotional and neutral images was at about 180 ms for faces and 220 ms for IAPS images.

----

insert Figure 2 about here

3.3 SSVEP amplitudes

Given the shorter latency for the first emotional ERP effect for faces compared to IAPS images, the question of interest is now, whether or not differential effects as a function of emotionality occurs earlier in the time course of SSVEP amplitudes with facial background images as well. This was indeed the case since we found the first time point with significantly greater reduction of SSVEP amplitudes for fearful compared to neutral faces at ~180 ms ($t_{17} = -1.75$; $p < 0.05$) compared to ~550 ms for IAPS images ($t_{17} = -1.75$; $p < 0.05$). The respective time courses are depicted in Figure 3.

----

insert Figure 3 about here

These differences cannot be due to differences in the baseline (i.e. during the presentation of the scrambled version), since we found no significant differences with running t-tests in the time window from 550 to 150 ms before picture onset. Note that due to the temporal resolution of the Gabor filter data points closer to the change to a concrete image would be contaminated by that change.

To further test for the “following emotional cue extraction hypothesis” we calculated the difference amplitudes between emotional and neutral stimuli (emotional minus neutral), respectively for faces and IAPS separately. A negative amplitude value signifies a greater SSVEP amplitude with a neutral background stimulus at the respective time point.

Figure 4 shows the difference curves representing the effects of emotion for faces and IAPS separately. The direct comparison of these difference curves showed a significant difference from 398 ms (t_{17} = -1.74; $p < 0.05$) up to 445 ms (t_{17} = -1.75; $p < 0.05$). Thus, SSVEP amplitudes were significantly reduced for fearful versus neutral faces earlier than for unpleasant versus neutral IAPS (gray area in Figure 4). Notably, both curves show an early downward deflection after the onset of the background image (better visible in Figure 3). However, this deflection only continues and reaches a long-lasting plateau for faces, rebounding upwards for IAPS images. The effect in IAPS begins a descent to a plateau somewhat later, at approximately 450 ms.

----

insert Figure 4 about here

3.4 Target detection rates

Figure 5 depicts the time course of target detection rates for all four experimental conditions. Note that hits were calculated time-locked to target onsets. Therefore, peaks around time point zero and for all other windows correspond to responses that, on average, occurred about 600 ms later. We found significant main effects of Time ($F_{7,119}$=20.73, $p<0.001$, η_g=0.24), Stimulus ($F_{1,17}$=3.13, $p<0.05$, η_g=0.004) and Emotion ($F_{1,17}$=9.15, $p<0.01$, η_g=0.009), with a higher detection rate when faces compared to IAPS images were presented in the background, and higher detection rates for neutral compared to emotional stimuli. Further, there were significant two-way interactions Stimulus x Emotion ($F_{1,17}$=19.25, $p<0.001$, η_g=0.03), Stimulus x Time ($F_{7,119}$=11.51, $p<0.001$, η_g=0.06), Emotion x Time ($F_{7,119}$=3.9, $p<0.001$, η_g=0.02). However, these effects were further qualified by the presence of a significant three-way interaction Time x Stimulus x Emotion ($F_{7,119}$=2.6, $p<0.05$, η_g=0.013) which suggested that target rate differed with respect to stimulus type and emotionality in a certain time period. We followed up this interaction by calculating separate ANOVAs for each bin of the factor Time. For the sake of brevity, we will only focus here on the most crucial comparisons and effects (see Table 1 for more detailed statistics for all the comparisons).

Before face/picture change there was no significant difference in target detection rates between emotional and neutral stimuli (See Table 1 and Figure 5). Similar to our previous studies, hit rates to coherent motion events that started shortly before the change to a concrete image (-267 to 0 ms) dropped considerably (main effect of Stimulus, see Table 1). Crucially, target rates differed between faces and IAPS with respect to emotionality in the time period from 1 to 800 ms after the image onset (consecutive time bins 4, 5 and 6 in the Table 1), which gave rise to the interaction between Emotion and Stimulus, as clearly visible in Figure 5).

Post-hoc tests revealed a significant reduction of target detection rates for unpleasant compared to neutral IAPS pictures in those three consecutive time bins ($bin_{1\text{-}267ms}$: p=0.01; $bin_{268\text{-}534ms}$: p=0.01; $bin_{535\text{-}800ms}$: p=0.002), as well as for neutral faces compared to neutral IAPS pictures ($bin_{1\text{-}267ms}$: p=0.049), and for unpleasant IAPS compared to fearful faces ($bin_{268\text{-}534ms:}$ p<0.001, other comparisons were not significant: p>0.06).

----

insert Table 1 and Figure 5 about here

3.5 Discussion

The experiment tested the hypothesis that the extraction of emotional cues precedes the reduction of SSVEP amplitude for emotional background stimuli relative to neutral stimuli. We found that the earliest differences between emotional and neutral faces in the SSVEP occurred at 180 ms, around the same time as the earliest differences observed in the ERPs (i.e. the N170). In contrast, the earliest significant difference between emotional and neutral IAPS images in the SSVEP occurred later, at about 550 ms, despite earlier differences in the ERPs to IAPS images. When directly comparing the time course of the emotional effects in the SSVEP between face and IAPS, we found that they differed significantly from 398 to 445 ms. Thus, emotional effects in the SSVEP were observed significantly earlier for faces than for IAPS. This significant period in-between the two times at which emotional effects were separately observed for faces and IAPS is consistent with a period of “ramping up” - once the emotional bias begins, it takes time to reach its maximum – which is malleable by stimulus type, and thus the specific challenges presented by each stimulus type. This is not consistent with an account in which a relatively time-invariant shifting process, as described in the introduction, is responsible for the slow-biasing of SSVEP amplitudes. If such a mechanism were responsible for the latency of the differential SSVEP amplitude effect as a function of valence, we predicted that the effect should occur at approximately the same time for faces as for IAPS. Instead, we would argue that emotional cue extraction lasts longer for IAPS (at the N1-EPN complex at around 200-300 ms) compared to faces (at the N170 at about 170 ms), and thus at least the greatest portion of the competition for processing resources between the background image and the foreground task does not occur until after this extraction.

Before further discussing the results of Experiment 1 in depth, we note that the latency of the SSVEP emotional effect with IAPS images was somewhat later (at about 550 ms) than in our other studies, which found emotional effects at around 400 ms with longer (Hindi Attar et al. 2010; Schönwald and Müller 2014) and 475 ms with a very short presentation (Müller et al. 2011). A critical difference between our previous studies and the present one are the adjustments to faces and IAPS images we made here to match their low-level visual properties as thoroughly as possible. These adjustments in luminance, contrast, and spatial frequency across all experimental conditions resulted in an overall decrease of clarity of the stimulus material, particularly for IAPS images. Thus, our low-level image property adjustments might have resulted in a lack of clarity that delayed the initial stages of content processing and lead to the later affective modulation of SSVEP amplitudes. The fact that the N1 peak was also delayed by about 30 ms compared to our previous study (Schönwald and Müller 2014) supports that idea. In order to test for that possibility of an delayed content processing for IAPS images due to our adjustments, we conducted a second experiment in which we compared the luminance, contrast, and spatial frequency matched images we used here with the identical but unadjusted, non-matched stimuli using an emotion categorization task.

4. Methods of Experiment 2

4.1 Subjects

We recruited twenty-five volunteers (19 female) to participate in the study after giving their informed written consent. All observers had normal or corrected to normal vision and their ages ranged from 18 to 43 years (M=26, SD=5.68). One subject was excluded due to a history of epileptic seizures in the past. This resulted in overall 24 participants. All of them were compensated for their participation time (6 Euros/hour).

4.2 Stimuli and task

We used the identical grayscale images as in Experiment 1. In order to test the influence of the matching procedure for low-level properties we additionally presented faces as well as IAPS images in their non-matched original grayscale version (see Figure 6 for examples). In order to avoid recognition or priming effects, each participant saw an individual IAPS image or face only once in either its “matched" or “non-matched” (i.e. original) version. That required two groups of participants. Group one saw 50% of IAPS/faces in their matched and the other 50% in their original version. In group two the order was reversed, i.e. group two saw the matched versions of group one in their original version and vice versa. The use of the matched or non-matched version of each stimulus was counterbalanced across participants, so that, overall, each image appeared as “matched” or “non-matched” an equal number of times.

As reported above, the adjusted images had a mean luminance of 0.44 (representative of the global luminance), SD=0.0004 and a mean contrast=0.096 (representative of RMS contrast), SD=0.0002, with no differences in luminance or contrast between conditions (see Methods, Stimuli and task for Experiment 1). The identical but non-matched stimuli differed in luminance (mean luminance = 0.4 of grayscale value, SD = 0.12, $F_{3,296}=6.44$, $p<0.001$, partial $\eta^2=0.06$) and contrast (mean = 0.22; SD = 0.06; $F_{3,296}=42.75$, $p<0.001$, partial $\eta^2=0.302$) between conditions. Overall the design comprised of three factors: Thus, there were 3 factors in the experiment: Matching (matched vs. non-matched), Stimulus (IAPS vs. face), Emotion (emotional vs. neutral).

----

insert Figure 6 about here

Subjects performed an emotion categorization task and were instructed to report as fast as possible whether the content of the presented image/face was negative or neutral. They were asked to press the letter “C” on the keyboard whenever a fearful face or an unpleasant picture was shown or the letter “B” whenever a neutral face/picture was presented on the screen. Participants were instructed to use only one finger for the response and maintained their gaze at fixation throughout the entire trial. In a comfortable seated position stimuli were presented centrally on a computer screen at a distance of 80 cm (13.3 x 10.71 degrees of visual angle) as in the Experiment 1. A centrally located white cross served as fixation point during the entire experiment on a 19 inch computer screen. The screen was set at a resolution of 800x600 pixels and 16 bits per pixel color mode and a screen refresh rate of 60 Hz.

Each image was presented for 1000 ms with an inter-trial interval that varied between 1150 and 1450 ms. The experiment consisted of 4 blocks with 60 trials each, resulting in 240 trials in total. Participants were given a short break by the end of each block.

4.3 Analysis

Only reaction times to correct responses from 250 to 1000 ms were included in the analysis. A repeated-measures analysis of variance (ANOVA) with the within-subjects factors of Stimulus (IAPS, Face), Emotion (negative, neutral) and Matching (matched, non-matched) was calculated. To further analyze differences between conditions with respect to stimulus type, we conducted post-hoc t-tests using the Bonferroni-Holm correction for multiple comparisons where necessary.

5. Results

Overall, subjects responded correctly for 94.07% of all images. Statistical analysis of reaction times revealed generally slower responses to neutral than to emotional stimuli (main effect of Emotion type: $F_{1,23}=23.38$, $p<0.001$, $\eta_g=0.048$), as well as overall longer response times to IAPS pictures than facial expressions (main effect Stimulus type: $F_{1,23}=49.61$, $p<0.001$, $\eta_g=0.19$) and generally increased RTs towards matched compared to non-matched stimuli (main effect of Matching: $F_{1,23}=21.32$, $p<0.001$, $\eta_g=0.015$). However, the interpretation of these main effects is qualified by the presence of significant two-way interactions of Emotion type x Stimulus type ($F_{1,23}=4.81$, $p<0.05$, $\eta_g=0.014$) as well as Stimulus type x Matching ($F_{1,23}=12.174$, $p<0.01$, $\eta_g=0.008$).

Post-hoc t-tests on the Emotion type x Stimulus type interaction showed that differences in RTs between fearful and neutral face were not significant ($p=0.08$), whereas all the other comparisons were ($ps<0.001$). Specifically, there was a significant difference in RTs between unpleasant and neutral IAPS pictures, with slower reaction times for neutral than unpleasant IAPS pictures ($p<0.001$), slower responses for both neutral IAPS than neutral faces ($p<0.001$) and for unpleasant IAPS than fearful faces ($p<0.001$), see Table 2. Additionally, reaction times were slower for unpleasant IAPS than neutral faces, and faster for fearful faces than neutral IAPS ($ps<0.001$).

Of particular interest here are the post-hoc t-tests on the interaction between Stimulus Type and Matching (Figure 7, see also Table 2). These tests revealed that response times between matched and non-matched faces did not differ significantly ($p=0.37$), but importantly, subjects responded significantly slower to matched compared to non-matched IAPS images ($p<0.001$). The comparisons also showed faster reaction time for matched faces than matched IAPS ($p<0.001$) as well as for non-matched faces than non-matched IAPS ($p<0.001$), indicating that responses were faster to overall faces than IAPS pictures irrespective of whether they were matched or not.

----

insert Table 2 and Figure 7 about here

Importantly, neither the interaction Emotion type x Matching ($F_{1,23}=0.08$, $p=0.79$, $\eta_g=0.00008$), nor the interaction Emotion type x Stimulus type x Matching was significant ($F_{1,23}=3.38$, $p=0.08$, $\eta_g=0.003$), suggesting that matching stimuli in their low-level properties had no effect on the emotional processing of any type of the presented stimulus (face or IAPS).

5.1 Discussion

Experiment 2 tested whether matching the spatial frequency energy, luminance and contrast of the faces and IAPS pictures used in Experiment 1 leads to slower reaction times in an emotion categorization task than when using images with identical emotional content but not matched for low-level properties. We found that reaction times were significantly slower in response to matched compared to non-matched IAPS pictures, but there was no effect of matching for face stimuli. These reaction time differences support the idea of a delay in initial stages of stimulus processing for low-level adjusted IAPS images, reflected in the delay of the N1-peak compared to their original version. Importantly, although participants tended to be slower in categorizing neutral than unpleasant IAPS, low-level stimulus properties adjustments had no influence on the emotional cue extraction. That was true for faces as well as for IAPS images. Not surprisingly, reaction times to faces were also significantly faster compared to IAPS images, probably due to the difference in complexity between faces and IAPS.

6. General Discussion

Competition for processing resources between objects in our visual environment is a key neural mechanism (Desimone 1998; Kastner and Ungerleider 2001). Emotional stimuli have the capacity to bias processing resources in their favor, resulting in preferred processing due to their intrinsic biological relevance (Bradley et al. 2003; Lang and Davis 2006). In our standard distraction paradigm, which allows investigation of the temporal dynamics of competition for processing resources in early visual cortex, we previously found attention-emotion competition to occur at a latency of about 400 ms with complex emotional IAPS images (Hindi Attar et al. 2010; Müller et al. 2008; Schönwald and Müller 2014). We hypothesized that competition for processing resources in early visual cortex is preceded by the extraction of the emotional cue, a process linked to the N1-EPN ERP complex observed for IAPS images, which has a typical latency of around 200 ms or even longer (Weinberg and Hajcak 2010; Wiens and Syrjänen 2013). Although the results of our recent study suggested this hypothesis (Schönwald and Müller 2014), we could not exclude an alternative explanation. Rather than following emotional cue extraction, the onset of competitive interactions could follow a relatively time-invariant shifting process that might consist of two steps that follow each other but most likely overlap in time. First, there is capture of attention due to "objecthood", likely expressed by the first SSVEP amplitude deflection. Second, there may be a feature-based shift towards individual features of the objects. By using two different types of stimuli, faces and IAPS images, for which emotional cues are extracted quickly and more slowly respectively, we were here able to test whether competitive interactions would occur earlier for faces than IAPS (consistent with the "following emotional cue" hypothesis) or at approximately the same time (consistent with the time-invariant shifting hypothesis). The innovation of our design is to allow for concurrent analysis for the VEP elicited by the background stimuli and the time course of competition for processing resources in early visual cortex providing significant complementary information on neural dynamics.

As in our previous studies (Hindi Attar et al. 2010; Müller et al. 2008; Schönwald and Müller 2014), we found that SSVEP amplitude reductions were significantly greater for affective distractors than for their neutral counterparts. The earliest time at which an SSVEP amplitude emotional modulation was observed was approximately 180 ms for faces but not until about 550 ms for IAPS pictures, although the emotional effect per se did not differ between faces and IAPS until 398 ms, once the emotional effect in faces had peaked and before the emotional effect in IAPS had fully developed (see Figures 3 and 4). Our results with regard to the VEP elicited by the background stimulus nicely replicated a number of previous findings. We found that the N1-EPN and the N170 elicited a more negative deflection towards emotionally charged than towards neutral IAPS and faces respectively. Contrary to some reports (Pourtois, Thut et al. 2005; Pourtois, Dan et al. 2005), we found no modulation of the P1 component as a function of facial expression. In concert, recent evidence shows that P1 component modulation by face-like properties can be fully accounted for by low-level visual cues (i.e. differences in facial features; Rossion and Caharel 2011). Thus, the first ERP component to show emotional sensitivity for facial expressions was the N170 (peaking at about 175 ms) for faces and the N1 (at about 220 ms) with the following EPN modulation for IAPS images. Interestingly, latency of the N1 peak as well as the onset of the differential SSVEP amplitude effect was much later compared to our previous study (Schönwald and Müller 2014).

Thus, we found that the SSVEP effects followed or at least began no sooner than the ERP effects. The relative delay between the emotional effect in the SSVEP in response to faces compared to the response to IAPs images indicates that the SSVEP effect follows emotional cue extraction rather than being due to a time-invariant shifting process. An analysis of responses to faces alone would have rendered these possibilities difficult to disentangle with the effect of emotional facial expression on the N170 and the affective modulation of the SSVEP for faces occurring at about the same latency (~180 ms). Nevertheless, the temporal smoothing inherent in the SSVEP analysis method (approximately ±110 ms here) may underestimate effect onsets to a degree, bringing them slightly further forward in time. Thus, the physiological response measured using the SSVEP most likely evolves somewhat later than the N170 effect. Furthermore, the SSVEP emotional effect continues to increase after it becomes significant. Nevertheless, the rapidity with which SSVEP changes occur with background faces at least suggests that the emotional cue extraction for faces is quickly completed and can rapidly bring about top-down attentional biases. For IAPS images, this stage takes longer, with the clear delay between the EPN emotional effect for IAPS images and the emotional effect in the SSVEP strongly consistent with attentional biasing necessarily following emotional cue extraction. If emotional cue extraction were not required, then SSVEP effects should not be relatively delayed for faces or IAPS.

Experiment 2 was set out to test whether these shifts in latency might be due to the matching of physical stimulus characteristics (i.e. brightness, contrast, spatial frequency) between the two image categories. We found significantly prolonged reaction times in the emotion categorization task for matched relative to unmatched IAPS images but no differences between matched and unmatched faces. Crucially, the matching process did not interact with emotion, suggesting that participants were able to efficiently extract emotion as well from matched as from unmatched images. In addition, this provides further evidence that the emotional cue extraction from the IAPS images was later than that for faces. In fact, our matching may have increased the disparity in timing between the emotional cue extraction for the two categories, since it delayed emotion categorization more for IAPS images than faces. Thus, this result provides more support to our "following emotional cue" hypothesis than a time-invariant shifting account. Taken together, results from Experiment 2 complement electrophysiological findings from the first experiment with longer latencies for the N1 peak and the SSVEP amplitude effect by suggesting that the image adjustments procedure resulted in prolonged early identification processes for IAPS images as a main factor that did not affect valence.

A second hint that competitive interactions in early visual cortex follows emotional cue extraction of the background image was provided by our recent study in which we presented background IAPS images for 200 ms, followed by an immediate mask (Müller et al. 2011). The presentation duration of 200 ms was selected in order to allow for emotional content categorization but no further elaborative processing (Larson et al. 2005; Schupp, Junghöfer et al. 2004). We found significantly reduced SSVEP amplitudes for unpleasant compared to pleasant images at a latency of about 475 ms. Interestingly, that was 275 ms after the disappearance and masking of the background image. As in our previous studies, the onset of the background image caused an immediate drop in SSVEP amplitude that did not differ as a function of valence, and is likely to be related to a capture of attention due to “objecthood”. If a second feature-based shifting process were required, then that shifting process would have taken place in the absence of the concrete image for the greatest portion of its time. Our results rather suggest that once the emotional content of a background distractor is extracted it seems to trigger competitive interactions in early visual cortex irrespective of whether or not the emotional distractor is still visible. Thus, our results strongly support the “following emotional cue extraction” hypothesis.

Behavioral data from Experiment 1 also yielded interesting results. Overall, we found the typical pattern of a decrease in target detection accuracy from picture onset lasting until 800 ms after picture onset. This decrease was also greater overall for IAPS pictures than for faces, and for emotional relative to neutral stimuli. Notably, however, emotion and stimulus type interacted; the emotional distraction effect was significant for IAPS but not for faces. The fact that we did not find behavioral costs as a function of emotional facial expression replicates a finding of one of our previous studies. In that yet unpublished study we presented angry, fearful, happy, and neutral facial expressions from the FACES Life Span Database of Facial Expressions (Ebner et al. 2010) with posers of Caucasian origin with a distraction paradigm as used in the current study and tested 17 participants. As with the present study, we found a significant drop in target detection rates with the onset of the background face, but no statistically significant differences with respect to valence (see Figure 8).

----

insert Figure 8 about here

Nevertheless, we can only cautiously interpret the null finding with regard to emotional facial expressions. One reason might be related to the fact that facial expressions are predominantly considered to be involved in emotion recognition (i.e. recognize anger being expressed on the face of the observer) with little capacity for provoking emotional reactions in the viewer. In comparison, more complex IAPS scenes are seen to induce stronger emotional reactions due to their higher arousal than images with faces (cf. Bradley et al. 2003; Britton et al. 2006). This argument is consistent with the ratings we obtained for facial expressions in the present study, which were rated as less arousing and not as unpleasant as IAPS scenes. Furthermore, some previous behavioral evidence suggested that task-irrelevant emotional faces, although capable of involuntarily "capturing" sensory processing resources in visual cortex, do not exhibit a significant impact on behavior (Fenker et al. 2010; Kujawa et al. 2012). Nevertheless, the capture of attention was clearly visible in the initial drop in SSVEP amplitudes as well as in target detection rates.

Finally, it might well be the case that the temporal resolution of our behavioral measure was not fine enough to show valence effects for faces due to the relatively rapid emotional cue extraction. It is well known that faces rapidly capture attention (Bannerman et al. 2010; Brosch et al. 2008; Tsuchiya et al. 2009). When subjects were required to make a saccadic response to a picture of a face or to another distractor-picture (such as animals or vehicles) that were simultaneously flashed in the left and right visual hemifield, the earliest saccades were directed to the competing face-stimulus with a latency between 100 to 110 ms and about 150 ms on average (Crouzet et al. 2010). That was contrasted by mean saccadic response times for vehicles with about 190 ms. Given that face detection in saccadic choice task studies (cf. Crouzet et al. 2010; Kirchner and Thorpe 2006) is much faster and emotional cue extraction of facial expressions might be finished with the N170 component (Gilles Pourtois, personal communication), future studies that use a better temporal resolution might shed light on our somewhat surprising findings with respect to facial distractors. It is also worth noting that the effect of emotional facial expression on the N170 remains yet unsettled in the literature, with considerable amount of studies reported no modulation of the N170 with facial affect (Dennis and Chen 2007; Holmes et al. 2005; Schupp, Öhman et al. 2004). Recent evidence by Rellecke and colleagues suggest that some methodological factors such as electrode reference might provide a potential explanation to so far conflicting evidence with regard to N170 modulations with emotion (Rellecke et al. 2013). Nevertheless, our electrophysiological findings clearly show greater competitive interactions in early visual cortex for fearful compared to neutral facial expressions. With low-level image properties (luminance and contrast) being equalized between the two face categories, the modulation of the N170 component remained present.

6.1 Conclusion

This is the first study that investigated the time course of competition between distracting task-irrelevant emotional background IAPS images and facial expressions during the performance of a visual foreground task in a within subjects design. We found that the reduction in SSVEP amplitudes towards unpleasant IAPS pictures was preceded by the N1-EPN modulation. For faces, emotional content extraction occurred at the level of N170 and was also followed by the SSVEP affective modulation. The emotional distraction effect in the SSVEP occurred earlier for faces than IAPS pictures, in line with the temporal relationship between the emotional cue extraction process indexed by the N170 and N1-EPN ERP components. The present findings, thus, indicate that the attentional resource competition bias in early visual areas follows emotional content discrimination for both stimulus types, faces and more complex IAPS images, and occurs later when emotional cue extraction occurs later. Our data do not allow us to unambiguously draw conclusions on the neural mechanism, i.e. whether the impact on SSVEP amplitude is due to re-entrant feedback mechanisms from higher cortical areas (Keil et al. 2009). Nevertheless, our results are most consistent with a framework in which there are reciprocal and/or competitive interactions between cortical systems that guide top-down (endogenous) modulation and the circuits that guide emotional attention in early visual cortex, as suggested recently (Pourtois et al. 2013).

Acknowledgments

The authors thank Renate Zahn, Maike Salazar-Kämpf and Anna Pfau for their kind assistance in data acquisition and Dr. Sonja A. Kotz for her valuable assistance in the interpretation of results. The study was supported by Deutsche Forschungsgemeinschaft, Graduate school Function of Attention in Cognition and grant MU972/22-1.

Conflict of interest: none declared.

Address of the corresponding author: Matthias M. Müller, Institute of Psychology, University of Leipzig, Neumarkt 9-19, 04109 Leipzig, Germany. Email: m.mueller@rz.uni-leipzig.de.

References

Andersen SK, Müller MM. 2010. Behavioral performance follows the time course of neural facilitation and suppression during cued shifts of feature-selective attention. Proc Natl Acad Sci USA. 107:13878-13882.

Anderson AK, Christoff K, Panitz D, De Rosa E, Gabrieli J. 2003. Neural correlates of the automatic processing of threat facial signals. J Neurosci. 23:5627-5633.

Ashley V, Vuilleumier P, Swick D. 2004. Time course and specificity of event-related potentials to emotional expressions. Neuroreport. 15:211-216.

Baguley T. 2012a. Serious stats: a guide to advanced statistics for the behavioral sciences. Houndmills, Basingstoke, Hampshire (England); New York (NY): Palgrave Macmillan.

Baguley T. 2012b. Calculating and graphing within-subject confidence intervals for ANOVA. Behav Res Methods. 44:158-175.

Bakeman R. 2005. Recommended effect size statistics for repeated measures designs. Behav Res Methods. 37:379-384.

Bannerman R, Milders M, Sahraie A. 2010. Attentional bias to brief threat-related stimuli revealed by saccadic eye movements. Emotion. 10:733–738.

Batty M, Taylor M. 2003. Early processing of the six basic facial emotional expressions. Brain Res Cogn Brain Res. 17:613-620.

Bentin S, Allison T, Puce A, Perez E, McCarthy G. 1996. Electrophysiological studies of face perception in humans. J Cogn Neurosci. 8:551-565.

Bentin S, Deouell L. 2000. Structural encoding and identification in face processing: ERP evidence for separate mechanisms. Cogn Neuropsychol. 17:35-55.

Blau V, Maurer U, Tottenham N, McCandliss B. 2007. The face-specific N170 component is modulated by emotional facial expression. Behav Brain Funct. 3:7.

Bradley MM, Lang PJ. 1994. Measuring emotion: the Self-Assessment Manikin and the semantic differential. J Behav Ther Exp Psychiatry. 25:49-59.

Bradley MM, Sabatinelli D, Lang PJ, Fitzsimmons JR, King W, Desai P. 2003. Activation of the visual cortex in motivated attention. Behav Neurosci. 117:369-380.

Britton JC, Taylor SF, Sudheimer KD, Liberzon I. 2006. Facial expressions and complex IAPS pictures: common and differential networks. Neuroimage. 31:906-919.

Brosch T, Sander D, Pourtois G, Scherer KR. 2008. Beyond fear: rapid spatial orienting toward positive emotional stimuli. Psychol Sci. 19:362-370.

Crouzet SM, Kirchner H, Thorpe SJ. 2010. Fast saccades toward faces: face detection in just 100 ms. J Vis. 10:1–17.

Dan-Glauser E, Scherer K. 2011. The Geneva affective picture database (GAPED): a new 730-picture database focusing on valence and normative significance. Behav Res Methods. 43:468-477.

Delorme A, Makeig S. 2004. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. J Neurosci Methods. 134:9-21.

Dennis TA, Chen CC. 2007. Neurophysiological mechanisms in the emotional modulation of attention: The interplay between threat sensitivity and attentional control. Biol Psychol. 76:1-10.

Desimone R, Duncan J. 1995. Neural mechanisms of selective visual attention. Annu Rev Neurosci. 18:193-222.

Desimone R. 1998. Visual attention mediated by biased competition in extrastriate visual cortex. Philos Trans R Soc Lond B Biol Sci. 353:1245-1255.

Di Russo F, Pitzalis S, Aprile T, Spitoni G, Patria F, Stella A, Spinelli D, Hillyard SA. 2007. Spatiotemporal analysis of the cortical sources of the steady-state visual evoked potential. Hum Brain Mapp. 28:323-334.

Ebner NC, Riediger M, Lindenberger U. 2010. FACES - a database of facial expressions in young, middle-aged, and older women and men: development and validation. Behav Res Methods. 42:351-362.

Eimer M, Holmes A. 2007. Event-related brain potential correlates of emotional face processing. Neuropsychologia. 45:15-31.

Fenker DB, Heipertz D, Boehler CN, Schoenfeld MA, Noesselt T, Heinze H-J, Duezel E, Hopf J-M. 2010. Mandatory processing of irrelevant fearful face features in visual search. J Cogn Neurosci. 22:2926-2938.

Grill-Spector K, Malach R. 2004. The human visual cortex. Annu Rev Neurosci. 27:649-677.

Hindi Attar C, Andersen SK, Müller MM. 2010. Time course of affective bias in visual attention: Convergent evidence from steady-state visual evoked potentials and behavioral data. Neuroimage. 53:1326-1333.

Holmes A, Winston JS, Eimer M. 2005. The role of spatial frequency information for ERP components sensitive to faces and emotional facial expression. Cognitive Brain Research. 25:508-520.

Holmes A, Kiss M, Eimer M. 2006. Attention modulates the processing of emotional expression triggered by foveal faces. Neurosci Lett. 394:48-52.

Junghöfer M, Elbert T, Tucker DM, Rockstroh B. 2000. Statistical control of artifacts in dense array EEG/MEG studies. Psychophysiology. 37:523-532.

Junghöfer M, Bradley MM, Elbert TR, Lang PJ. 2001. Fleeting images: a new look at early emotion discrimination. Psychophysiology. 38:175-178.

Kastner S, De Weerd P, Desimone R, Ungerleider LG. 1998. Mechanisms of directed attention in the human extrastriate cortex as revealed by functional MRI. Science. 282:108-111.

Kastner S, Ungerleider LG. 2001. The neural basis of biased competition in human visual cortex. Neuropsychologia. 39:1263-1276.

Keil A, Moratti S, Sabatinelli D, Bradley MM, Lang PJ. 2005. Additive effects of emotional content and spatial selective attention on electrocortical facilitation. Cereb Cortex. 15:1187-1197.

Keil A, Sabatinelli D, Ding M, Lang PJ, Ihssen N, Heim S. 2009. Re-entrant projections modulate visual cortex in affective perception: evidence from Granger causality analysis. Hum Brain Mapp. 30:532-540.

Kirchner H, Thorpe SJ. 2006. Ultra-rapid object detection with saccadic eye movements: visual processing speed revisited. Vision Res. 46:1762-1776.

Krombholz A, Schaefer F, Boucsein W. 2007. Modification of N170 by different emotional expression of schematic faces. Biol Psychol. 76:156-162.

Kujawa A, Klein DN, Hajcak G. 2012. Electrocortical reactivity to emotional images and faces in middle childhood to early adolescence. Dev Cogn Neurosci. 2:458-467.

Lang P, Bradley M, Cuthbert BN. 2008. International affective picture system (IAPS): affective ratings of pictures and instruction manual. Technical Report A-6. Gainesville (FL): University of Florida.

Lang PJ, Davis M. 2006. Emotion, motivation, and the brain: reflex foundations in animal and human research. Prog Brain Res. 156:3-29.

Larson CL, Ruffalo D, Nietert JY, Davidson RJ. 2005. Stability of emotion-modulated startle during short and long picture presentation. Psychophysiology. 42:604-610.

Morgan ST. 1996. Selective attention to stimulus location modulates the steady-state visual evoked potential. Proc Natl Acad Sci USA. 93:4770-4774.

Müller MM, Picton TW, Valdes-Sosa P, Riera J, Teder-Sälejärvi WA, Hillyard SA. 1998. Effects of spatial selective attention on the steady-state visual evoked potential in the 20-28 Hz range. Brain Res Cogn Brain Res. 6:249-261.

Müller MM, Andersen S, Trujillo NJ, Valdes-Sosa P, Malinowski P, Hillyard SA. 2006. Feature-selective attention enhances color signals in early visual areas of the human brain. Proc Natl Acad Sci USA. 103:14250-14254.

Müller MM, Andersen SK, Keil A. 2008. Time course of competition for visual processing resources between emotional pictures and foreground task. Cereb Cortex. 18:1892-1899.

Müller MM, Andersen SK, Hindi Attar C. 2011. Attentional bias to briefly presented emotional distractors follows a slow time course in visual cortex. J Neurosci. 31:15914-15918.

Öhman A, Flykt A, Esteves F. 2001. Emotion drives attention: detecting the snake in the grass. J Exp Psychol Gen. 130:466-478.

Okon-Singer H, Lichtenstein-Vidne L, Cohen N. 2013. Dynamic modulation of emotional processing. Biol Psychol. 92:480-491.

Olejnik S, Algina J. 2003. Generalized eta and omega squared statistics: measures of effect size for some common research designs. Psychol Methods. 8:434-447.

Olofsson JK, Nordin S, Sequeira H, Polich J. 2008. Affective picture processing: an integrative review of ERP findings. Biol Psychol. 77:247-265.

Pessoa L, McKenna M, Gutierrez E, Ungerleider LG. 2002. Neural processing of emotional faces requires attention. Proc Natl Acad Sci USA. 99:11458-11463.

Petersen SE, Posner MI. 2012. The attention system of the human brain: 20 years after. Annu Rev Neurosci. 35:73-89.

Pourtois G, Dan ES, 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. Hum Brain Mapp. 26:65-79.

Pourtois G, Thut G, Grave de Peralta R, Michel C, Vuilleumier P. 2005. Two electrophysiological stages of spatial orienting towards fearful faces: early temporo-parietal activation preceding gain control in extrastriate visual cortex. Neuroimage. 26:149-163.

Pourtois G, Schettino A, Vuilleumier P. 2013. Brain mechanisms for emotional influences on perception and attention: what is magic and what is not. Biol Psychol. 92:492-512.

R Core Team. 2012. R: A language and environment for statistical computing. Vienna (Austria): R Foundation for Statistical Computing.

Regan D. 1989. Human brain electrophysiology: Evoked potentials and evoked magnetic fields in science and medicine. New York (NY): Elsevier.

Rellecke J, Sommer W, Schacht A. 2012. Does processing of emotional facial expressions depend on intention? Time-resolved evidence from event-related brain potentials. Biol Psychol. 90:23-32.

Rellecke J, Sommer W, Schacht A. 2013. Emotion Effects on the N170: A Question of Reference? Brain Topogr. 26:62-71.

Rossion B, Caharel S. 2011. ERP evidence for the speed of face categorization in the human brain: Disentangling the contribution of low-level visual cues from face perception. Vision Res. 51:1297-1311.

Rousselet G, Pernet C, Bennett P, Sekuler A. 2008. Parametric study of EEG sensitivity to phase noise during face processing. BMC Neurosci. 9:98.

Schendan HE, Lucia LC. 2010. Object-sensitive activity reflects earlier perceptual and later cognitive processing of visual objects between 95 and 500 ms. Brain Res. 1329:124-141.

Schönwald LI, Müller MM. 2014. Slow biasing of processing resources in early visual cortex is preceded by emotional cue extraction in emotion–attention competition. Hum Brain Mapp. 35:1477-1490.

Schreij D, Owens C, Theeuwes J. 2008. Abrupt onsets capture attention independent of top-down control settings. Percept Psychophys. 70:208-218.

Schreij D, Theeuwes J, Olivers CL. 2010. Abrupt onsets capture attention independent of top-down control settings II: additivity is no evidence for filtering. Atten Percept Psychophys. 72:672-682.

Schupp HT, Junghöfer M, Weike AI, Hamm AO. 2003. Attention and emotion: an ERP analysis of facilitated emotional stimulus processing. Neuroreport. 14:1107-1110.

Schupp HT, Junghöfer M, Weike AI, Hamm AO. 2004. The selective processing of briefly presented affective pictures: an ERP analysis. Psychophysiology. 41:441-449.

Schupp HT, Öhman A, Junghöfer M, Weike AI, Stockburger J, Hamm AO. 2004. The facilitated processing of threatening faces: An ERP Analysis. Emotion. 4:189-200.

Schupp HT, Flaisch T, Stockburger J, Junghöfer M, Anders S, Ende G, Junghofer M, Kissler J, Wildgruber D. 2006. Emotion and attention: event-related brain potential studies. Prog Brain Res. 156:31-51.

Schupp HT, Stockburger J, Bublatzky F, Junghöfer M, Weike AI, Hamm AO. 2008. The selective processing of emotional visual stimuli while detecting auditory targets: an ERP analysis. Brain Res. 1230:168-176.

Thorpe S, Fize D, Marlot C. 1996. Speed of processing in the human visual system. Nature. 381:520-522.

Tsuchiya N, Moradi F, Felsen C, Yamazaki M, Adolphs R. 2009. Intact rapid detection of fearful faces in the absence of the amygdala. Nat Neurosci. 12:1224-1225.

Vuilleumier P, Armony JL, Driver J, Dolan RJ. 2001. Effects of attention and emotion on face processing in the human brain: an event-related fMRI study. Neuron. 30:829-841.

Weinberg A, Hajcak G. 2010. Beyond good and evil: the time-course of neural activity elicited by specific picture content. Emotion. 10:767-782.

Widmann A. 2006. Firfilt EEGLAB Plugin. Version 1.5.1. Leipzig (Germany): University of Leipzig.

Wiens S, Syrjänen E. 2013. Directed attention reduces processing of emotional distracters irrespective of valence and arousal level. Biol Psychol. 94:44-54.

Willenbockel V, Sadr J, Fiset D, Horne G, Gosselin F, Tanaka J. 2010. The SHINE toolbox for controlling low-level image properties. Behav Res Methods. 42:671-684.

Yantis S, Jonides J. 1984. Abrupt visual onsets and selective attention: evidence from visual search. J Exp Psychol Hum Percept Perform. 10:601-621.

Yantis S, Jonides J. 1996. Attentional capture by abrupt onsets: new perceptual objects or visual masking? J Exp Psychol Hum Percept Perform. 22:1505-1513.

Appendix

IAPS numbers of neutral pictures:

1350, 1645, 1908, 1945, 2036, 2039, 2191, 2235, 2272, 2273, 2377, 2400, 2445, 2485, 2525, 2580, 2593, 2594, 2597, 2745, 5120, 5130, 5201, 5395, 5455, 5500, 5535, 5635, 7033, 7036, 7037, 7041, 7044, 7081, 7130, 7136, 7140, 7160, 7161, 7165, 7179, 7217, 7300, 7491, 7495, 7504, 7512, 7513, 7546, 7547, 7550, 7560, 7595, 7632, 7700, 7710, 8162, 8250, 8325 and 9260.

IAPS numbers of unpleasant pictures:

1050, 1111, 1120, 1202, 1300, 1304, 1525, 2717, 2730, 2981, 3001, 3015, 3016, 3019, 3064, 3103, 3110, 3140, 3150, 3190, 3191, 3195, 3212, 3213, 3250, 3261, 3400, 6190, 6230, 6263, 6300, 6415, 6570, 8480, 9102, 9140, 9181, 9183, 9185, 9187, 9301, 9322, 9342, 9400, 9405, 9420, 9471, 9480, 9495, 9500, 9561, 9570, 9571, 9596, 9623, 9635, 9810, 9901, 9902 and 9940.

Tables

Time window (ms)	**-800/-534**	**-533/-268ms**	**-267/0**	**1/267**	**268/534**	**535/800**	**801/1067**	**1068/1334**
Stimulus	ns	Ns	$F_{1,17}$=17.0 $p<0.001$ η_g=0.09	Ns	$F_{1,17}$=31.7 $p<0.001$ η_g=0.18	$F_{1,17}$=13.6 $p<0.01$ η_g=0.08	$F_{1,17}$=4.5 $p<0.05$ η_g=0.04	ns
Emotion	ns	Ns	ns	$F_{1,17}$=8.9 $p<0.01$ η_g =0.03	$F_{1,17}$=5.3 $p<0.05$ η_g =0.03	$F_{1,17}$=12.8 $p<0.01$ η_g =0.12	ns	ns
Stimulus*Emotion	ns	Ns	ns	$F_{1,17}$=5.8 $p<0.05$ η_g =0.02	$F_{1,17}$=16.3 $p<0.001$ η_g =0.13	$F_{1,17}$=13.2 $p<0.01$ η_g =0.08	ns	ns

Table 1. Statistical values for significant effects for each time bin resulting from the repeated measures ANOVA comprising the factors of Stimulus (face vs. IAPS) and Emotion (emotional vs. neutral) and Time (1-8 bins). Note: ns = non-significant.

Matched				**Non-matched**			
IAPS		**Face**		**IAPS**		**Face**	
Unpleasant	**Neutral**	**Fearful**	**Neutral**	**Unpleasant**	**Neutral**	**Fearful**	**Neutral**
666±16	705±16	593±11	615±16	627±14	684±17	594±14	603±13

Table 2. Mean (±standard error of the mean) for reaction times (ms) for matched vs. non-matched fearful and neutral faces as well as unpleasant and neutral IAPS pictures in the emotion categorization task.

Captions to figures

Figure 1: Experimental design and iso-contour voltage map of the grand average SSVEP. (**A**) A trial started with the presentation of the scrambled view of the image. After a variable time interval it changed to the intact view of either a neutral or a fearful face or a neutral or unpleasant IAPS picture. Superimposed red squares constituted the foreground task. (**B**) Topographical distribution of the grand average 15 Hz SSVEP amplitude across all subjects and conditions for the time period of 1 sec before and 2 sec after the image change, with a clear peak around electrode Oz. Note: Example for unpleasant and neutral pictures is taken from GAPED database (Dan-Glauser and Scherer 2011).

Figure 2. ERP waveforms on the left and iso-contour voltage maps on the right for N170, N1 and EPN components. (**A**) Grand mean N170 waveform averaged across all subjects for fearful (dashed line) and neutral faces (solid line) across electrodes P9 and P10. On the right is the topographical distribution for N170 in the time window between 160 and 200 ms. (**B**) Grand mean N1 waveform averaged for unpleasant (dashed line) and neutral pictures (solid line) across the cluster of 12 electrodes labeled at the topographical distribution of the difference value (unpleasant minus neutral) between 200 and 240 ms. (**C**) Grand mean EPN waveform averaged for unpleasant (dashed line) and neutral pictures (solid line) across identical electrodes as for the N1. The topographical distribution is shown for the difference value (unpleasant minus neutral) between 275 and 425 ms.

Figure 3. Gabor filtered SSVEP time course for emotional face or IAPS background images. **(A)** Grand average of SSVEP amplitudes for neutral (solid line) and fearful background faces (dotted line) at electrode Oz. (**B**) Grand average of SSVEP amplitudes for neutral (solid line) and unpleasant IAPS pictures (dotted line) at electrode Oz. Note: The vertical gray line at time point zero in both panels indicates the onset of the face/picture (change from the scrambled version of an image to the intact one). The vertical dotted lines with the star shows the onset of significant SSVEP differential amplitude effects between emotional and neutral background images.

Figure 4. Difference curves for SSVEP amplitudes. Grand average SSVEP amplitude differences for fearful minus neutral faces (dotted line) and unpleasant minus neutral IAPS pictures (solid line) at electrode Oz. Thin vertical line at time point zero indicates the onset of the face/picture (change from

the scrambled version of an image to the intact one). Gray box marks the time window of a significant difference between emotional effect in faces and IAPS (from ~ 400 to ~450 ms). For orientation the dashed and solid vertical line with the star show the time point of the significant differences in SSVEP amplitudes for emotional vs. neutral faces or IAPS at 180 and 550 ms, respectively.

Figure 5. Behavioral data. Time course of target detection rates in percent for 8 consecutive time windows (3 before and 5 after the image change) averaged across all participants when neutral faces (black solid line), fearful faces (black dotted line) or neutral IAPS (gray solid line) or unpleasant IAPS pictures (gray dotted line) were presented in the background. The change from a scrambled to a concrete image occurred at the end of time bin -267/0.

Figure 6. Examples of the stimulus material of Experiment 2. **B** and **C** show an example of matched images for a fearful face and for a neutral picture as used in Experiment 1. Note: Example for a neutral picture is taken from GAPED database (Dan-Glauser and Scherer 2011). **A** and **D** are the respective non-matched, original versions of the same images.

Figure 7. Interaction between Stimulus Type and Matching conditions. Mean reaction times in ms for non-matched vs. matched faces (gray) as well as non-matched vs. matched IAPS (black). The 95% CIs correspond to the within-subject variability after between-subjects differences were removed (Baguley 2012b).

Figure 8. Behavioral data of facial expressions experiment. Time course of target detection rates in percent for 8 consecutive time windows (3 before and 5 after the image change) averaged across all participants when neutral (gray solid line), fearful (gray dashed line), angry (black dashed line) or happy (black solid line) faces were presented in the background. The change from a scrambled to a concrete face occurred at the end of time bin -267/0.