

Rapid sensory gain with emotional distracters precedes attentional deployment from a foreground task

Valeria Bekhtereva^a, Matt Craddock^b, Christopher Gundlach^a, Matthias M. Müller^{a,*}

^a Institute of Psychology, University of Leipzig, Leipzig, Germany

^b Department of Psychology, University of Lincoln, Lincoln, UK

ARTICLE INFO

Keywords:

Visual attention
Emotional scenes
Visual distraction paradigm
RSVP
SSVEP

ABSTRACT

The steady-state visual evoked potential (SSVEP), an electrophysiological marker of attentional resource allocation, has recently been demonstrated to serve as a neural signature of emotional content extraction from a rapid serial visual presentation (RSVP). SSVEP amplitude was reduced for streams of emotional relative to neutral scenes passively viewed at 6 Hz (~167 ms per image), but it was enhanced for emotional relative to neutral scenes when viewed as 4 Hz RSVP (250 ms per image). Here, we investigated whether these seemingly contradictory observations may be related to different dynamics in the allocation of attentional resources as a consequence of stimulation frequency. To this end, we advanced our distraction paradigm by presenting a visual foreground task consisting of randomly moving squares flickering at 15 Hz superimposed on task-irrelevant RSVP streams shown at 6 or 4 Hz, which could unpredictably switch from neutral to unpleasant content during the trial or remained neutral. Critically, our findings demonstrate that affective distractors captured attentional resources more strongly than their neutral counterparts, irrespective of whether they were presented at 6 or 4 Hz rate. Moreover, the emotion-dependent attentional deployment from the foreground task was temporally preceded by sustained sensory facilitation in response to emotional background images. Together, present findings provide evidence for rapid sustained visual facilitation but a rather slow attentional bias in favor of emotional distractors in early visual areas.

1. Introduction

One of the most extensively studied research topics in the field of attention and emotion concerns the extent to which distracting emotional cues can involuntarily capture attention (Carrette, 2014). The emotional salience of visual images has been frequently linked to facilitated processing, or sensory gain in response to emotional cues at the expense of other concurrent stimuli that compete for limited processing resources (Desimone, 1998; Desimone and Duncan, 1995; Keil et al., 2009). What remains less understood, however, is the time-scale of potential trade-off effects in the competition for visual processing resources between task-relevant stimuli and distracting emotional information, in particular, when multiple affective distractors are presented over a prolonged period of time.

Over the last decade, research utilizing steady-state visual evoked potentials (SSVEP) has made a significant contribution to our understanding of the neural time-scale of visual distraction effects with emotional stimuli (Deweese et al., 2014; Keil et al., 2002, 2012; Müller

et al., 2008; Trauer et al., 2012). The SSVEP is a brain response to a periodically presented stimulus with the same fundamental frequency response as the driving stimulus, and has neural sources primarily in early visual cortical areas (Di Russo et al., 2007; Müller et al., 2006; Regan, 1989; Wieser et al., 2016). Notably, SSVEP amplitude is significantly increased when a stimulus is attended compared to when it is ignored (Andersen et al., 2010; Morgan, 1996; Müller et al., 2006; Norcia et al., 2015; Vialatte et al., 2010). Most importantly, one can directly and continuously quantify the allocation of processing resources devoted to multiple simultaneously presented stimuli by tagging each stimulus with its unique frequency and obtaining its distinct SSVEP response. This property renders the SSVEP a powerful tool for studying attention-emotion interactions at early visual areas through dissociating and independently tracking the neural activity evoked by concurrently presented task-related stimuli and affective distractors.

In this regard, we have conducted a number of studies utilizing our distraction paradigm to investigate the time-course of competitive interactions between a foreground attentional task and either neutral or

* Corresponding author. Institute of Psychology University of Leipzig, Neumarkt 9-19, 04109, Leipzig, Germany.

E-mail addresses: valeriya.bekhtereva@uni-leipzig.de (V. Bekhtereva), m.mueller@rz.uni-leipzig.de (M.M. Müller).

emotional distracting pictures from the International Affective Picture Set (IAPS, Lang et al., 2008) that were always task-irrelevant (Bekhtereva et al., 2015; Deweese et al., 2016; Hindi Attar et al. 2010, Andersen et al., 2010; Müller et al., 2011; Müller et al., 2008; Müller and Gundlach, 2017; Schönwald and Müller, 2014). In these studies, participants were instructed to detect and respond to an occasional period of a coherent motion event of a flickering random dot kinematogram (RDKs) in the foreground while ignoring the distracter-image that could unpredictably appear during the trial in the background. Critically, the RDK flashes at a specific frequency, which elicits an SSVEP response at that frequency. SSVEP amplitude time-course analyses uncovered a significantly greater decrease in the SSVEP following the presentation of an emotional relative to a neutral background distracter-image at ~400–500 ms after picture onset that lasted for several hundred of milliseconds (Hindi Attar et al., 2010; Müller et al., 2008; Müller and Gundlach, 2017; Schönwald and Müller, 2014), indicating a greater withdrawal of attentional resources away from the visual task toward the emotional compared to the neutral picture during that time window.

Most recently, we have significantly extended those findings by demonstrating that the SSVEP can serve as a direct signature of emotional content extraction. In a series of experiments employing a rapid serial visual presentation paradigm (RSVP), neutral or unpleasant images were displayed in rapid succession to elicit the SSVEP (Bekhtereva & Müller, 2015, 2017b; Bekhtereva et al., 2018). Critically, a new scene of a particular valence category was displayed at each cycle of the presentation to assess whether SSVEP modulation can be driven by the repeated extraction of affective content from an image (cf. Bekhtereva et al., 2018). To this end, every image was shown at a relatively low-frequency range of either 3, 4, 6, 8.57 Hz (from ~117 to 333 ms per image) or at a fast 15 Hz rate (67 ms per image). The rationale for these presentation times was based on previous evidence indicating that emotional content extraction can be maintained with picture presentation rates up to ~12 Hz, as indexed by the robust amplitude modulation of early negativities (N1 and EPN visual evoked potentials, Peyk et al., 2009) elicited in response to the presentation of emotional relative to neutral images. As anticipated, we found no discernable differences in SSVEPs at 15 Hz RSVP. By contrast, SSVEP amplitudes were reliably modulated when a neutral relative to a negative RSVP was presented at 3–8.67 Hz rates, thereby providing the first evidence that the SSVEP modulation during affective image viewing is contingent upon the extraction of emotional features from each image.

However, the most striking observation from these experiments was related to the differential direction of emotional modulation at 6 Hz as opposed to other presentation rates. We observed a robust reduction in SSVEP amplitudes when streams of emotional relative to neutral images were viewed at 6 Hz (Bekhtereva and Müller, 2015, 2017), at odds with earlier studies demonstrating an SSVEP enhancement in response to visual objects that convey emotional information (Deweese et al., 2016; Gruss et al., 2012; Keil et al., 2003; Keil and Ihssen, 2004; Wieser and Keil, 2011, 2014; Wieser et al., 2012) and in contrast with the robust increase in SSVEP amplitude for emotional scenes at all lower presentation rates, with particularly well-pronounced effects at 4 Hz. Moreover, the reversed SSVEP amplitude modulation was also replicated with 6.67 Hz RSVP and generalized to pleasant images, strongly indicating that the inverted modulation was not driven by a specific emotional valence (cf. Bekhtereva et al., 2018).

How could such opposite patterns of the observed SSVEP amplitude modulation with emotional images be reconciled? Our latest findings strongly suggested that the SSVEP attenuation with affective content at ~6 Hz rate may be accounted for by a linear superposition of ERPs (Capilla et al., 2011). More specifically, systematic differences in transient ERPs in response to emotional as compared with neutral images in an RSVP stream may have led to similar systematic differences in the

SSVEP, which, in turn, may have increased or decreased the power at the driving frequency, thus resulting in the reversed amplitude modulation pattern (for more details see Bekhtereva et al., 2018). However, an alternative possibility remains such as neutral images received preferential processing and attracted more attentional resources during picture viewing at ~6 Hz. Consequently, this may have resulted in a distinct, facilitated perceptual processing, driving larger SSVEP amplitudes to neutral as compared to emotional images.

The objectives of the present experiment were thus two-fold. On the one hand, we aimed to confirm whether a 6 Hz RSVP of neutral or emotional IAPS images would demonstrate evidence of differential image processing to a 4 Hz RSVP stream, with more attentional resources attracted to neutral images during the 6 Hz presentation. On the other hand, we wanted to examine whether neural competition for visual processing resources between a primary task and emotional distracters exhibits time-point-by-time-point reciprocity. In particular, we investigated whether or not the withdrawal of attentional resources from the task could simultaneously be accompanied by reciprocal attentional pull or sensory gain for the distracter-images. To this end, we adapted our distraction paradigm by presenting a background RSVP stream at either 4 or 6 Hz in which all images were either of neutral valence, or would change from neutral to unpleasant images at some point during the trial. As with our previous studies, the foreground task was defined by an RDK flickering at 15 Hz. The crucial methodological advantage of the current design over earlier studies is that presenting multiple scenes with an unexpected change from neutral to emotional valence within an RSVP does not trigger an onset ERP in response to the content change and, as a result, does not disrupt the ongoing SSVEP time-courses elicited by the task and emotional distracters. This is achieved based on each presentation cycle being defined by the onset of a new image as opposed to the sudden visual onset of a single, identical neutral or emotional image presented flickering for several seconds, as was the case in our previous studies (Deweese et al., 2016; Keil et al., 2009).

The new design allowed us to concurrently analyze the time-course of SSVEP amplitudes elicited by the 15 Hz foreground task and the 6 or 4 Hz RSVP background stream. Of interest here was whether we would find time-invariant instantaneous shifting of attentional resources away from the foreground task to the emotional background images as would be predicted by practically all attentional models. We hypothesized that if larger SSVEPs with neutral compared to emotional images at 6 Hz genuinely reflect an allocation of a more considerable amount of attentional resources toward neutral content, then this should be mirrored in a higher amplitude reduction (higher withdrawal of processing resources) of the 15 Hz foreground SSVEP when neutral as opposed to emotional RSVPs are shown at 6 Hz. Alternatively, if emotional scenes are stronger competitors for visual processing resources relative to their neutral counterparts, as repeatedly follows from our findings from the distraction paradigm cited above, a significant reduction in the 15 Hz SSVEP should be observed when emotional relative to neutral RSVPs are shown at both 4 and 6 Hz rates. With regard to the temporal dynamics of the shift of attentional resources, we know from our previous studies that the task-related SSVEP amplitude reduction occurs relatively late, between 400 and 550 ms after the presentation of a distracting emotional relative to a neutral scene, suggesting that the withdrawal of attentional resources follows the higher-order process of emotional cue identification as was reflected in the modulation of N1 and EPN components at ~200–400 ms (Bekhtereva et al., 2015; Schönwald and Müller, 2014). If this were the case, we would expect a biphasic process of, first, neural facilitation of the distracting emotional RSVP stream, followed by a reduction of the task-related SSVEP amplitude as a consequence of the shift of attentional resources towards the task-irrelevant background stream, with a larger attentional shift for emotional compared to neutral images forming the RSVP stream.

2. Materials and methods

2.1. Participants

Based on the smallest effect size obtained in our original study, which had a similar RSVP design (Bekhtereva and Müller, 2015) with a 6 Hz presentation rate ($\eta_g^2 = 0.19$ or $\eta_p^2 = 0.429$), an a priori power analysis suggested that a sample size of 21 volunteers should be sufficient to achieve power of 0.95. The power analysis was done using G*Power software (Faul et al., 2007). Given the current study's primary purpose of replicating and extending our previous findings, a somewhat larger sample was obtained to account for the more complex experimental design and to ensure adequate statistical power. Thirty-three participants (26 female, seven male) with a mean age of 23 years (standard deviation [SD] = 5.07) with normal or corrected to normal visual acuity took part in the study. Three participants were excluded from the final sample: the EEG recording of one participant was heavily contaminated with eye blinks, while the data of the other two subjects could not be fully saved due to technical issues. Thus, thirty healthy volunteers (24 female, six male) with a mean age of 24 years ([SD] = 5.2) were included in the final sample.

Prior to the experimental recording, participants received information about the study goals and gave written informed consent. All subjects received either class credits or monetary compensation for their participation (8 € per hour). The study was approved by the ethical review board of the University of Leipzig and conformed to the Code of Ethics of the World Medical Association.

2.2. Stimuli

Task-relevant stimuli constituted an array of 115 randomly distributed moving yellow squares ($0.6^\circ \times 0.6^\circ$ of visual angle) superimposed on unpleasant or neutral images. Picture size subtended $10.36^\circ \times 8.22^\circ$ of visual angle at a viewing distance of 80 cm.

A total of 60 neutral and 60 unpleasant color picture scenes were selected from the IAPS (Lang et al., 2008) and from the Emotional Picture Set databases (EmoPicS, Wessa et al., 2010). This set of pictures was used in our previous series of experiments utilizing a similar RSVP stimulation protocol (Bekhtereva et al., 2018). Unpleasant and neutral scenes differed on normative ratings of valence (unpleasant: $M = 2.55$ (SD = 0.70); neutral: $M = 5.08$ (SD = 0.54); $t_{111.46} = -22.13$, CI 95% = [-2.75–2.3], $p < 0.001$, $d = -4.04$) and arousal (unpleasant: $M = 6.06$ (SD = 0.74); neutral: $M = 3.40$ (SD = 0.75); $t_{117.99} = 19.5$, CI 95% = [2.39–2.93], $p < 0.001$, $d = 3.56$). Images were resized to 419×335 pixels and pre-processed with MATLAB image processing toolbox together with SHINE image toolbox (Willenbockel et al., 2010) to ascertain equal mean luminance and mean RMS contrast between experimental image categories. To this end, the mean (representative of the global luminance) and standard deviation (representative of root-mean-squared (RMS) contrast) of the luminance distribution of each picture scene were calculated on the intensity of pixels in the range between 0 and 1 (normalized RGB values, with minimum 0 = black and maximum 1 = white). A two-sample Welch t -test for unpleasant and neutral image sets showed no statistically significant differences for mean luminance ($t_{116.6} = 0.13$; mean difference = 0.000004 CI 95% = [-0.00006–0.00007]; $d = 0.02$; $p = 0.89$) or contrast composition ($t_{91.149} = 0.33$; mean difference = 0.00002 CI 95% = [-0.0001–0.0001]; $d = 0.06$; $p = 0.74$) between two conditions. Additionally, we calculated luminance against the screen background that was between 24 and 56 cd/m^2 .

Furthermore, subjective and objective ratings of picture complexity (obtained from Andreas Keil, University of Florida, Gainesville) were compared between unpleasant and neutral pictures. The two-sample Welch t -test did not demonstrate any statistically significant differences between emotional and neutral categories for either subjective complexity ($t_{117.94} = 0.85$; mean difference = 0.26 CI 95% = [-0.34

0.86]; $d = 0.15$; $p = 0.39$) or JPEG complexity ($t_{116.46} = -0.36$; mean difference = -0.74 CI 95% = [-4.74–3.27]; $d = -0.07$; $p = 0.72$).

2.3. Experimental procedure

Stimuli were presented at the center of a 19-inch computer screen set at a resolution of 1024×768 pixels against a black background, 16 bits per pixel color mode, and 60 Hz monitor refresh rate. A centrally located yellow cross served as fixation during the experiment. Yellow squares were randomly distributed over centrally presented pictures and were shown at a flicker frequency of 15 Hz (displayed “on” for two frames and “off” for two frames), to record a steady-state visual evoked potential (SSVEP). Pictures were presented in the background as a rapid serial visual presentation (RSVP) stream: each image was shown for 15 or 10 frames of screen refresh (250 and ~ 167 ms, respectively) at 4 and 6 Hz. The flow and timing of image presentation were controlled with the Cogent toolbox for MATLAB (Cogent, www.vislab.ucl.ac.uk/Cogent/; The Mathworks, Inc, Natick, Massachusetts).

Each trial began with the presentation of the background RSVP stream of neutral pictures superimposed with flickering yellow squares. At a certain variable time point during a trial, the RSVP stream of neutral picture scenes either changed to a stream of unpleasant images or remained neutral (Fig. 1). Changes in emotional valence were jittered and occurred randomly and only once during a trial in an early (~ 667 ms for 6 Hz RSVP; 2000 ms for 4 Hz RSVP; $\sim 3.5\%$ of trials), middle (2000–3000 ms for 6 Hz RSVP; 3000–4000 ms for 4 Hz RSVP; $\sim 93\%$ of trials) or late (4000 ms for 6 Hz RSVP; 6000 ms for 4 Hz RSVP; $\sim 3.5\%$ of trials) time window after trial onset. This was performed in order to prevent any anticipation effects for change in emotional content. Trials with early and late changes in affective content served as “catch trials” and were excluded from analysis (total of 32 trials per participant).

Trial presentation period lasted 7500 ms for 4 Hz and 5000 ms for 6 Hz RSVP, resulting in 30 cycles of the respective presentation rate and thus 30 images shown per trial. Across the experiment, each neutral picture was presented ~ 168 times, and each unpleasant image was shown ~ 58 times (neutral images needed to be repeated more often, given that there were trials in which only neutral images were presented, without the change from neutral to unpleasant content). Overall, the experiment consisted of 452 trials subdivided into ten blocks (9 blocks with 45 trials each, and the last 10th block with 47 trials, respectively), with 420 trials that were included in the final data analysis (105 trials per condition). Each experimental block lasted ~ 5 min, and subjects were allowed to take a break between each block. Each image was presented in a randomized order (with the restriction that the presentation of two identical images in subsequent order never occurred). After each trial, a black screen with a white cross was presented for an interval varying randomly between 1150 and 1650 ms.

Every frame of 60 Hz screen refresh rate (16.667 ms) the squares moved in a random direction to the left, right, up, or down by 0.04° of visual angle. Targets were defined as events in which 45–60% of the squares moved coherently for four frames of the stimulation frequency of 15 Hz (i.e., 266.68 ms) in one of the four cardinal directions (up, down, left or right). During each trial between 0 and 4 coherent motion events (targets) could occur. In order to allow for analysis of the time-course of the behavioral data, targets were uniformly distributed across each cycle of 15 Hz stimulation (i.e., 66.67 ms), such that across the experiment three targets occurred in each time window of 66.67 ms in every experimental condition. The duration of the target event was 266.7 ms (i.e., four cycles of 15 Hz). Participants were instructed to detect targets as accurately and quickly as possible and to press “Space” on the keyboard while disregarding the background RSVPs as task-irrelevant. Halfway through the experimental session, participants switched the responding hand, and the starting hand was counterbalanced across all subjects. All conditions were presented randomized across trials. All participants performed a short training session (up to 4 blocks) on the recording day. For this practice session, we presented a different set of background

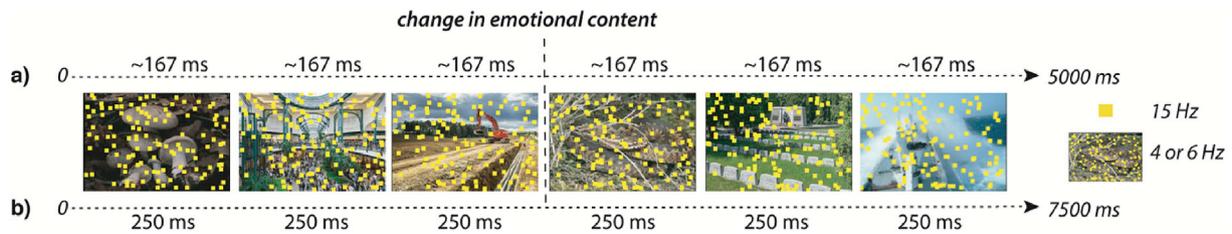


Fig. 1. A schematic example of a trial with a change from neutral to unpleasant content in the background. Each trial started with the presentation of flickering at 15 Hz yellow squares that were superimposed on an RSVP stream of neutral scenes, which after a variable time interval changed to an RSVP or unpleasant scenes (or remained neutral in other trials). Every image was shown for either ~ 167 ms (a) or 250 ms (b), corresponding to a 6 or 4 Hz RSVP rate, respectively. Examples of pictures for the illustration are obtained from <https://commons.wikimedia.org>.

images. Critically, to avoid ceiling effects of performance across participants, during the training session, the task difficulty was adjusted and defined in a stepwise procedure for each participant: the percentage of flickering squares that coherently move simultaneously in one cardinal direction varied between 45% and 60%. When participants achieved a stable target rate of at least $\sim 60\%$ with the respective percentage of coherently moving squares, the experimental EEG recording followed, and the established level of task difficulty during the practice session was used further in the EEG experiment.

Following the EEG session, participants viewed the pictures used in the experiment in randomized order and were asked to rate them on the dimensions of affective arousal and valence on the 9-point Self-Assessment-Manikin (SAM) scale (Bradley and Lang, 1994). For the picture rating procedure, we kept the exposure time of each image identical to that viewed in the EEG experiment. During the picture rating procedure, to maintain similar masking conditions to that in the EEG study, participants were shown an image from the experiment that was briefly flashed for either 250 ms (one cycle of 4 Hz presentation rate) or ~ 167 ms (one cycle of 6 Hz rate) and then immediately masked by its phase-scrambled (content-distorted) version displayed for the same duration (cf. Bekhtereva et al., 2018). Afterward, the SAM rating scale was displayed, asking to give a rating of arousal and valence for the respective picture by pressing a corresponding number on the keyboard. Overall, the entire set of experimental images was shown twice: for 250 ms in the first block and ~ 167 ms in the second block, with the order of blocks counterbalanced across all participants.

2.4. EEG-recording and analysis

Electroencephalographic activity (EEG) was recorded using a BioSemi ActiveTwo system (BioSemi, The Netherlands) at a sampling rate of 512 Hz from 64 Ag/AgCl scalp electrodes mounted in an elastic cap according to the international 10-20 system (Jasper, 1958). Two electrodes were used as reference and ground (CMS – “Common Mode Sense” and DRL – “Driven Right Leg”); for details see <http://www.biosemi.com/faq/cms&drl.htm>) during the recording. Lateral and vertical eye movements, as well as blinks, were monitored with a horizontal and vertical electrooculogram (EOGs) through four bipolar electrodes that were positioned above and below the right eye as well as on the outer canthi of each eye.

2.5. SSVEP analysis

Data preprocessing was performed with custom-built MATLAB scripts (The Mathworks, Natick, MA) and functions in EEGLAB (Delorme and Makeig, 2004). Epochs were extracted from 2700 ms before to 2700 ms after the image change for 4 Hz trials, and from 1800 ms before to 1800 ms after the content change for 6 Hz trials. All trials, with and without targets, entered the analysis. An automatic procedure for detecting epochs contaminated with artifacts was done using ‘Statistical Control of Artifacts in Dense Array EEG/MEG studies’ (SCADS, Junghöfer et al., 2000). All epochs were subsequently manually screened for

artifacts, particularly for non-stereotypical artifacts, and rejected if contaminated. Data was subsequently re-referenced to the average reference, and linear trends were removed. Following that, epochs were subjected to an independent component analysis (ICA), to remove sources of ocular and muscle artifacts (Delorme et al., 2012). The resulting ICA components were visually inspected, and those components identified as artefactual (i.e., showing typical topographies of eye artifacts or high frequency/line noise) were pruned from the data. In a final step, trials were averaged for each experimental condition.

The time-courses of SSVEP amplitudes were extracted by means of a Gabor filter centered at the stimulation frequencies of 4, 6, and 15 Hz, respectively (FWHM: ± 1.5 Hz, temporal resolution: ± 147.1 ms). Due to poor amplitude estimation by Gabor filter close to the edges of the data and in order to keep the epoch length for analysis equal across all experimental conditions, time windows from 1500 ms before and 1500 ms after the change in emotional content entered the statistical analysis.

A grand mean iso-contour voltage map based on the grand average across all time points, experimental conditions, and participants revealed greatest SSVEP amplitudes at a parieto-occipital cluster of eight electrodes: PO3, PO7, O1, Oz, Iz, PO4, PO8, and O2. Due to some variations in the topographical SSVEP amplitude distributions across individual subjects, in order to account for individual subject variability and improve the signal-to-noise ratio (SNR) of each stimulation frequency, one electrode with the largest amplitude was taken from that cluster for each subject and each frequency (the individual best electrode approach) for statistical analyses. Given that the electrode selection was based on an orthogonal contrast, it is not biased by the magnitude of amplitude differences between experimental conditions. The individual best electrode approach is commonly applied and used in the field (cf. Fuchs et al., 2008; Kaiser et al., 2005).

For the statistical analyses, running repeated-measures $2 \times 2 \times 2$ ANOVAs with within-subjects factors of Change in Content (Yes/No), Frequency Location (Foreground/Background), and Background Frequency (4 Hz/6 Hz) were performed on SSVEP data. To decompose any significant interactions, follow-up ANOVAs and post-hoc paired running t-tests were used. To control for Type I error, p-values were corrected with false discovery rate for all comparisons (FDR, see Benjamini and Hochberg, 1995).

2.6. Behavioral data and SAM rating analyses

Correct button presses within the time interval of 250 and 1000 ms following the onset of target motion were considered as hits, while later responses or no response were regarded as misses. Onsets of target events were distributed uniformly over the time window from 1000 ms before and 2000 ms after the onset of emotional content change of the background RSVPs at 4 and 6 Hz. Hits were calculated across time bins of 200 ms by averaging across three short bins of 66.67 ms, which resulted in 15-time windows of 200 ms each. Thus, the time interval from 1000 ms before up to 2000 ms after the change in emotional content (6 time bins before and 9 after) was analyzed using a $2 \times 2 \times 15$ repeated-

measures ANOVA with within-subjects factors of Change in Content (yes/no), Background Frequency (4 Hz vs. 6 Hz), and Time (bin 1-15). All calculations were performed with respect to the time bin in which the onset of the target event occurred. To further explore differences between conditions with respect to time bins, we conducted post-hoc t-tests using the Bonferroni-Holm correction for multiple comparisons.

Mean arousal and valence SAM ratings were analyzed by a 2×2 repeated measures ANOVA with the factors of Picture presentation time (250 ms vs. 167 ms) and Emotion (unpleasant vs. neutral).

2.7. Statistical analyses

Statistical analyses were conducted with R v3.4.1 (R Core Team, 2012). We used the following packages for data manipulation, visualization, and statistical analyses: *tidyr* v0.8.1, *magrittr* v1.5, *purrr* v0.2.5, *stats* v3.3.2, *ez* v4.4-0 (Lawrence, 2013), *lsr* v0.5 (Navarro, 2015), *Rmisc* v1.5, *ggplot2* v2.21 (Wickham, 2009). Timepoint-by-timepoint statistical tests, as well as EEG topographical scalp maps, were created using R package *eegUtils* 0.1.15.dev (Craddock; <https://github.com/craddm/eegUtils>).

Generalized eta-squared (η_g^2), as well as Cohen's d (d), are reported as measures of standardized effect size (Baguley, 2012; Bakeman, 2005; Olejnik and Algina, 2003). Furthermore, unstandardized effect sizes and their 95% confidence intervals are given for relevant comparisons.

2.8. Data and code availability statement

Data and analysis codes will be available upon direct request, given that we did not obtain a written consent from the participating subjects to upload data in the public domain. This strategy is also in line with guidelines of data protection of the German Research Foundation.

3. Results

3.1. SAM ratings

For valence ratings, the 2 (emotion) \times 2 (picture presentation time) within-subjects ANOVA indicated that there was a significant main effect of Emotion ($F_{1,29} = 305.38$; $p < .001$, $\eta_g^2 = 0.83$; mean difference = -2.52 , CI 95% = $[-2.81-2.22]$; $d = 3.19$), with unpleasant relative to neutral picture scenes rated as more negative. Main effect of picture presentation time was not statistically significant ($F_{1,29} = 0.92$; $p = 0.35$; $\eta_g^2 = 0.002$; mean difference = 0.05 , CI 95% = $[-0.05-0.14]$; $d = 0.17$). A significant interaction Emotion \times Picture presentation time ($F_{1,29} = 5.3$; $p = 0.028$, $\eta_g^2 = 0.006$; mean difference = -0.18 , CI 95% = $[-0.34-0.02]$; $d = 0.42$) and its follow-up pairwise comparisons

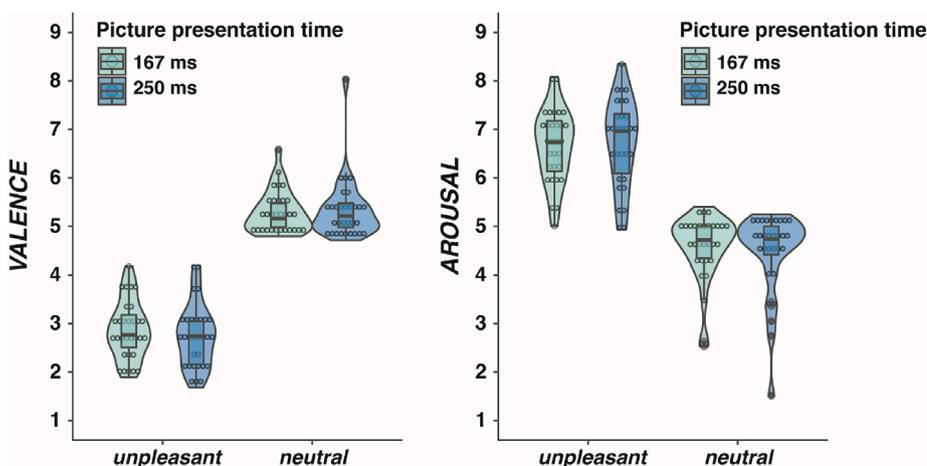


Fig. 2. Violin plots of subjective valence (A) and arousal (B) ratings for unpleasant and neutral scenes displayed as brief as ~ 167 ms (one cycle of 6 Hz) and 250 ms (one cycle of 4 Hz) for evaluations during post-experimental rating procedure. Picture presentation time had slightly opposite effects on valence ratings of images (A): while neutral images were rated the same between 167 and 250 ms conditions, unpleasant pictures were evaluated as slightly more negative, but this effect was minimal (0.13 mean differences). Note that subjective ratings of valence and arousal can receive values from 1 to 9 (1 corresponds to the most unpleasant valence and lowest arousal, and 9 reflects the most pleasant valence and greatest arousal). The width of the violin is determined by the kernel density estimate of the distribution of individual ratings (i.e., wider demonstrates that more participants rate within a given score). Boxplots are shown overlaid; individual data points are indicated by dark circles.

demonstrated that neutral images were evaluated similarly on valence, irrespective of their presentation time ($p = 0.53$; mean difference = -0.04 , CI 95% = $[-0.18-0.09]$; $d = 0.12$); however, emotionally negative pictures were rated as slightly more unpleasant relative to the same images when presented for a slightly longer time period of 250 ms compared to ~ 167 ms ($p = 0.03$; mean difference = 0.13 , CI 95% = $[0.03-0.24]$; $d = 0.45$), see Fig. 2.

For arousal ratings, there was neither a significant main effect of Picture presentation time ($F_{1,29} = 0.09$; $p = 0.76$, $\eta_g^2 = 0.00009$; mean difference = 0.02 , CI 95% = $[-0.08-0.11]$; $d = 0.57$), nor interaction Emotion \times Picture presentation time ($F_{1,29} = 2.5$; $p = 0.12$, $\eta_g^2 = 0.003$; mean difference = 0.17 , CI 95% = $[-0.05-0.39]$; $d = 0.29$). As expected, the main effect of Emotion was significant ($F_{1,29} = 115.14$; $p < .001$, $\eta_g^2 = 0.66$; mean difference = 2.17 , CI 95% = $[1.76-2.59]$; $d = 1.96$), with emotional relative to neutral scenes having overall higher arousal values (see Fig. 2).

3.2. SSVEP amplitudes

Running ANOVAs (F-tests) comprising within-subjects factors Change in Content (Yes/No), Frequency Location (Foreground/Background), and Background Frequency (4 Hz/6 Hz) were performed for each time point within the analysis time window from 1500 ms before and 1500 ms after the change in image valence. The results showed a significant three-way interaction Change in Content \times Frequency Location \times Background Frequency that began around 200 ms and lasted until 1500 ms ($F_s > 5.90$, $ps < 0.05$, $\eta_g^2 > 0.001$). To follow up this interaction, we computed two separate running ANOVAs for each level of the factor Frequency Location, and the significance threshold was set to $\alpha = 0.025$, respectively.

3.2.1. Background location (4 and 6 Hz RSVP image streams)

For background location, the results revealed that there was neither a significant main effect of Background Frequency ($F_s < 2.21$, $ps > 0.65$, $\eta_g^2 < 0.01$) nor a significant main effect of Change in Content ($F_s < 5.21$, $ps > 0.88$, $\eta_g^2 < 0.003$). Importantly, the interaction Change in Content \times Background Frequency was significant ($F_s > 7.55$, $ps < 0.02$, $\eta_g^2 > 0.004$) for the time period from ~ 217 ms to 1500 ms.

Follow-up paired running t-tests for simple main effects indicated that when background RSVPs were shown at 4 Hz, SSVEP amplitudes indicated a significant rise for emotional relative to neutral image streams between ~ 273 ms and 1500 ms ($t_{29} > -4.52$, $ps < 0.05$), as can be seen from Fig. 3A. By contrast, when background RSVP streams were presented at 6 Hz, SSVEP amplitudes dropped significantly for emotional relative to neutral image streams between ~ 225 ms and 1500 ms ($t_{29} > 2.43$, $ps < 0.05$), as depicted in Fig. 3B. Thus, the interaction was

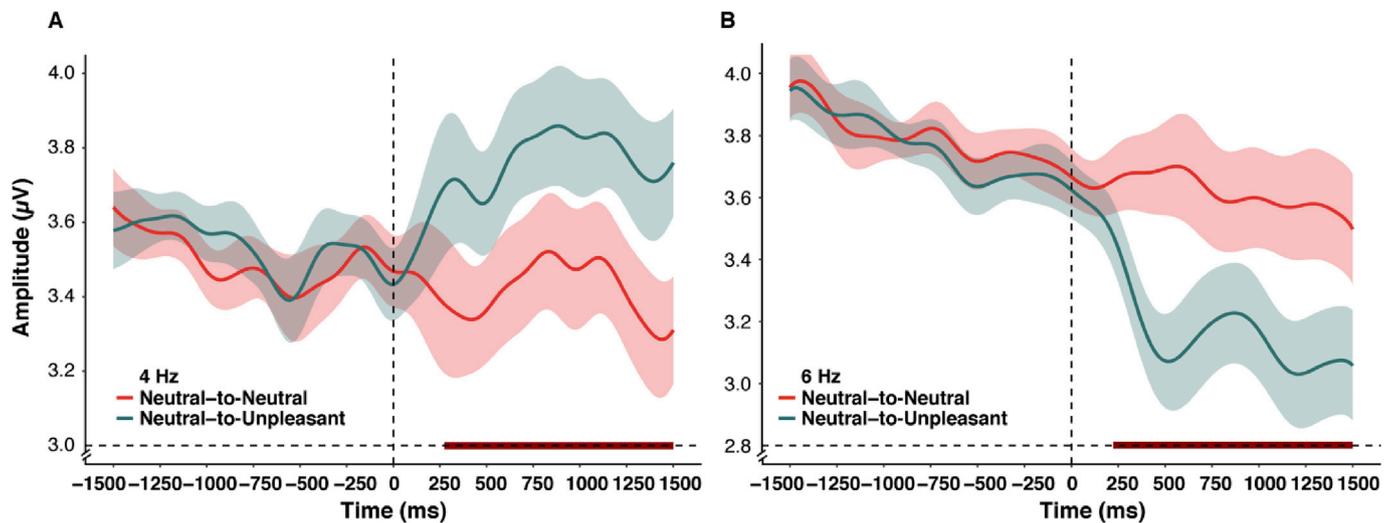


Fig. 3. Gabor-filtered SSVEP time-courses of background RSVP streams. Time point zero indicates a change in emotional content: from neutral to neutral (no change) or from neutral to unpleasant RSVP. A significant SSVEP differential amplitude (marked with a solid brown line; $p_s < 0.05$, FDR corrected) between neutral and emotional conditions develops at ~ 270 ms following a change in emotional content for 4 Hz RSVP (A) and at ~ 225 ms for 6 Hz RSVP (B). The shaded areas reflect 95% confidence intervals after between-subjects variability was removed using the Cousineau–Morey method (Morey, 2008).

driven by the differential direction of the SSVEP emotional effect for 4 Hz relative to 6 Hz SSVEP, while the onset of the emotional effects for both 4 and 6 Hz RSVPs occurred at about the same latency.

3.2.2. Foreground location (15 Hz task-related flicker)

For foreground location, there was no significant interaction Change in Content \times Background Frequency ($F_s < 5.8$, $p_s > 0.77$, $\eta_g^2 < 0.001$). The main effect of Background Frequency was significant ($F_s > 7.52$, $p_s < 0.02$, $\eta_g^2 < 0.03$), however, because this main effect is not relevant for our primary experimental question, we do not further focus here on its interpretation.

Notably, there was a significant main effect of Change in Content ($F_s > 8.4$, $p_s < 0.02$, $\eta_g^2 > 0.004$). Regardless of the RSVP frequency in the background, the foreground 15 Hz SSVEP demonstrated a significant drop in amplitude between ~ 472 ms and ~ 1316 ms ($F_s > 8.4$, $p_s < 0.02$, $\eta_g^2 > 0.004$) signifying a greater withdrawal of visual resources from the foreground task when unpleasant relative to neutral RSVPs were shown

in the background (see Fig. 4).

Importantly, in line with our hypothesis, we aimed to test for differences in the onset timing between emotion-driven attentional effects for the foreground task as compared with the valence-driven effects for background RSVPs. To this end, we calculated the difference amplitudes between “neutral-to-unpleasant” and “neutral-to-neutral” conditions and conducted two planned comparisons.

We directly tested the emotion-driven attentional effect of 15 Hz SSVEP (difference wave “neutral-to-unpleasant” minus “neutral-to-neutral” condition) against the valence-driven effect (difference wave “neutral-to-unpleasant” minus “neutral-to-neutral”) of the 4 Hz and 6 Hz background RSVP, respectively. As depicted in Fig. 5A–B, the results showed that differential effects between emotional SSVEP amplitude modulation at 15 Hz (foreground) vs. 4 Hz RSVP (background) were most pronounced from ~ 279 ms to 1500 ms ($t_{29} > -4.87$, $p_s < 0.05$). Similarly, when comparing time-courses of emotional modulation of the 15 Hz SSVEP signal at foreground with 6 Hz SSVEP at the background,

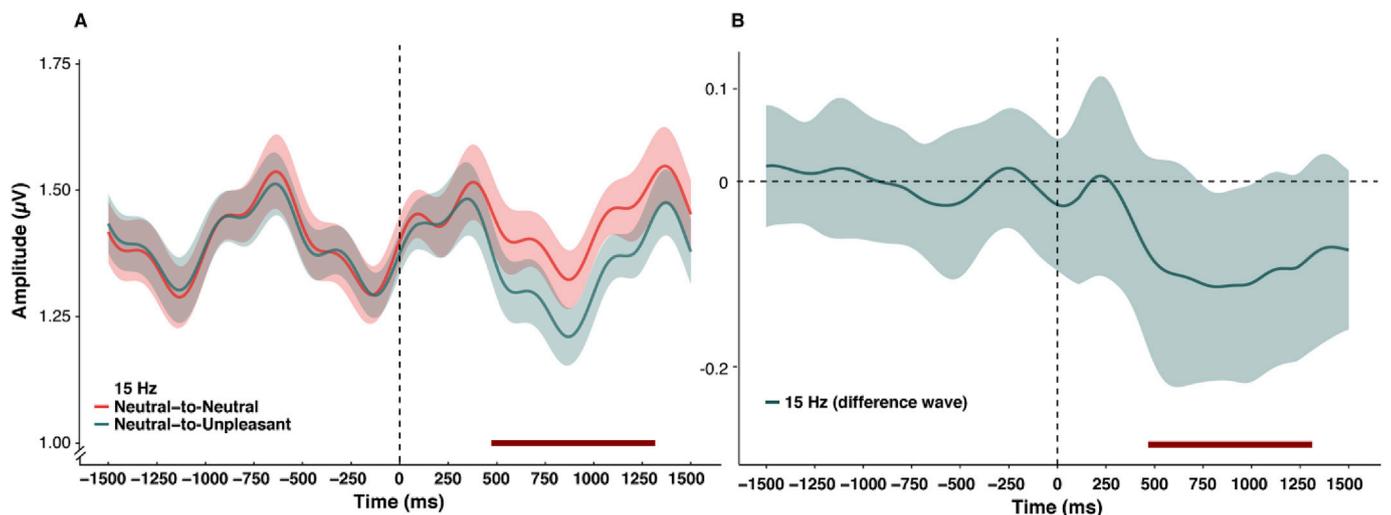


Fig. 4. A. Gabor-filtered 15 Hz SSVEP time-course elicited by flickering RDKs in the foreground while neutral or unpleasant RSVPs were presented in the background. Time point zero marks a change in emotional content in the background RSVP: from neutral to neutral (no change) or from neutral to unpleasant images. B. Gabor-filtered 15 Hz SSVEP time-course from (A) constructed as a difference wave (“neutral-to-emotional” minus “neutral-to-neutral” condition). The brown horizontal line in (A) and (B) indicates the significant differential SSVEP amplitude modulation between neutral and unpleasant conditions in the time interval from ~ 472 to ~ 1316 ms, irrespective of the background RSVP rate. The shaded areas represent 95% confidence intervals (Morey, 2008).

the results pinpointed statistically significant differences in the time windows from ~ 234 to 765 ms and ~ 920 –1500 ms ($t_{29} > 2.49$, $p < 0.05$). Thus, these direct comparisons supported the results of running t-tests for the foreground and background locations above and confirmed that the onsets of valence-driven emotional effects for both 4 and 6 Hz background RSVPs occurred earlier (by ~ 200 –250 ms) than that observed in the foreground with 15 Hz SSVEP emotion-driven attentional effects at ~ 470 ms after the change in emotional content. In addition, grand average SSVEP topographical maps for 4, 6, and 15 Hz SSVEP responses are depicted in Fig. 5C–E.

3.3. Target detection rates

We found significant main effects of Time ($F_{14,406} = 5.68$, $p < 0.001$, $\eta_g^2 = 0.028$), Emotion Change ($F_{1,29} = 4.32$, $p = 0.046$, $\eta_g^2 = 0.003$), Background Frequency ($F_{1,29} = 10.46$, $p = 0.003$, $\eta_g^2 = 0.009$), as well as a significant interaction Time x Background Frequency ($F_{14,406} = 1.85$, $p = 0.03$, $\eta_g^2 = 0.009$). However, the interpretation of these effects was qualified by the presence of the significant 3-way interaction: Time x Emotion Change x Background Frequency ($F_{14,406} = 1.94$, $p = 0.02$,

$\eta_g^2 = 0.008$). Similar to our previous study (Bekhtereva and Müller, 2017a), for the sake of brevity, we will consider below only the most relevant comparisons and effects: those that occurred in the time after the change in emotional content of the background RSVPs. For more details on the statistics, Table 1 contains additional information for all the comparisons and their unstandardized effect sizes together with 95% CIs where possible.

Overall, in the time windows ~ 1000 –1200 and ~ 1400 –2000 ms following the change in the emotional content of background RSVPs, detection rates for targets dropped significantly more for when 6 Hz relative to 4 Hz RSVP streams were shown, regardless of their emotional valence (Fig. 6). Although that decrease in performance lasted for several hundred milliseconds, it was rather small and constituted a drop of about seven percentage points on average. Furthermore, between ~ 1200 and 1400 ms after the change in emotional content, there was a significant interaction Emotion x Background Frequency ($F_{1,29} = 4.79$, $p = 0.037$, $\eta_g^2 = 0.02$). The follow-up pairwise comparisons, however, did not reveal any reliable significant differences between the RSVPs at 4 and 6 Hz as a function of emotional content (all $p > 0.12$, with maximal mean differences in hit rates between neutral and emotional RSVPs for 4 Hz at ~ 8.14

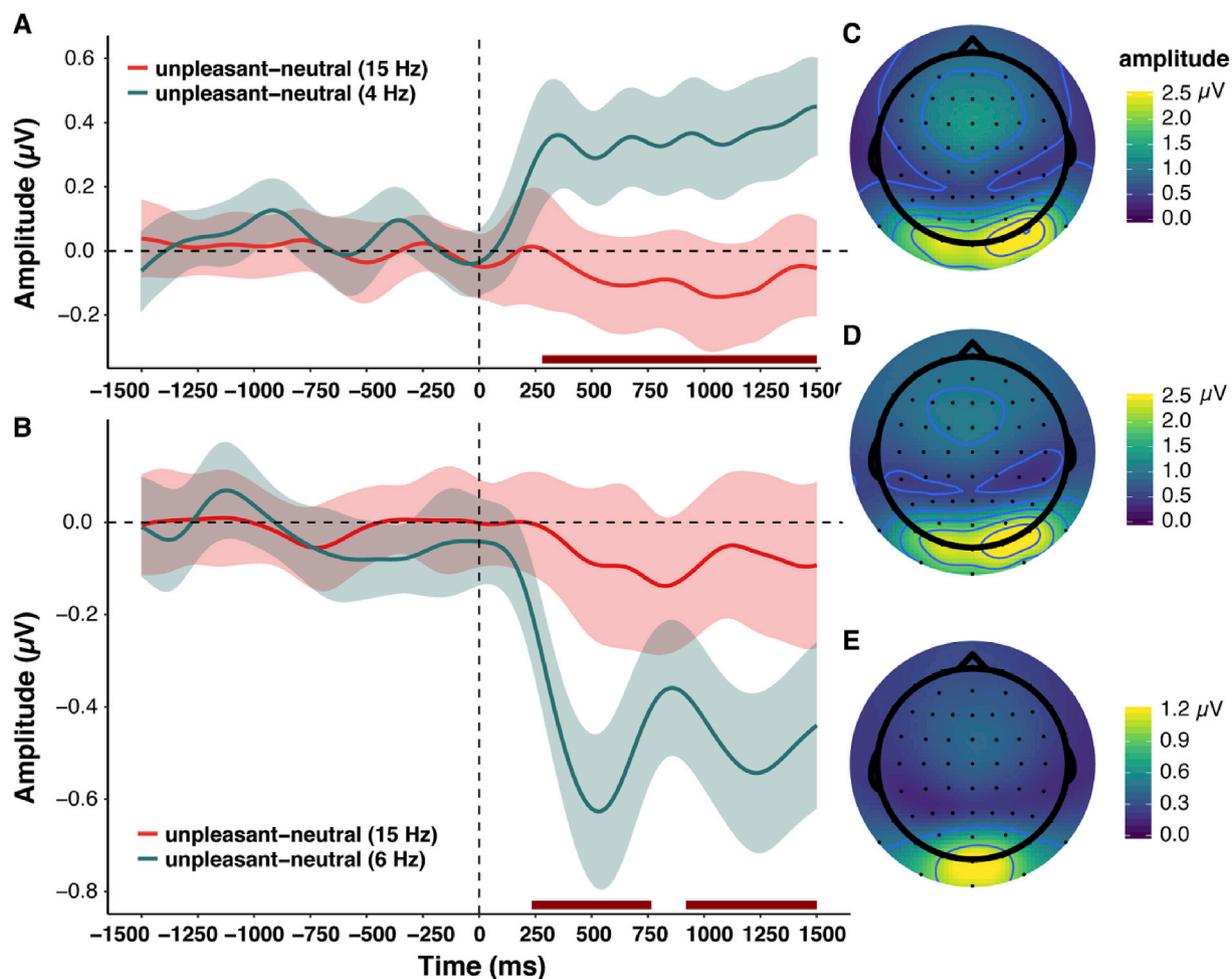


Fig. 5. SSVEP time-courses as difference waves (“neutral-to-emotional” minus “neutral-to-neutral” condition) and the topographical distributions of 4, 6, and 15 Hz SSVEP. The vertical line at time point zero shows the onset of the change in emotional content in the background RSVP stream. A negative amplitude value reflects larger SSVEP amplitudes with neutral RSVP streams, and a positive amplitude value signifies larger SSVEP amplitudes with unpleasant image streams. The shaded areas depict 95% confidence intervals (Morey, 2008). A. Differential effects between emotional SSVEP amplitude modulation at 15 Hz (red line) vs. 4 Hz RSVP (green line). The brown horizontal line on the x-axis indicates the significant ($p < .05$; FDR-corrected) time interval during which differences between the SSVEP emotion-driven attentional effects for foreground and valence-driven effects for background were most pronounced (from ~ 280 to ~ 1500 ms). B. Differential effects between emotional SSVEP amplitude modulation at 15 Hz (red line) vs. 6 Hz RSVP (green line). The brown horizontal line on the x-axis marks the significant ($p < .05$; FDR-corrected) time window during which differences between the SSVEP emotion-driven attentional effects for foreground and valence-driven effects for background were most pronounced (from ~ 235 to ~ 1500 ms). C–E. Topographical distribution of grand average for 4 Hz (C), 6 Hz (D), and 15 Hz (E) SSVEP amplitudes across the entire analysis window.

Table 1

Statistical values for each time bin from the repeated measures ANOVA comprising the factors of Emotion Change (from neutral to neutral vs. to unpleasant content) and Background Frequency (6 vs. 4 Hz). Note: n.s. = non-significant.

Time Window (ms)	Effect		
	Emotion	Frequency	Emotion x Frequency
-1000/-801	n.s. (p = 0.83)	n.s. (p = 0.06)	n.s. (p = 0.38)
	mean	mean	
	difference = 0.56% CI 95% = [-4.71 5.82]	difference = -3.52% CI 95% = [-7.19 0.16]	
-800/-601	n.s. (p = 0.68)	n.s. (p = 0.59)	n.s. (p = 0.64)
	mean	mean	
	difference = -1.11% CI 95% = [-6.65 4.43]	difference = -1.48% CI 95% = [-7.1 4.05]	
-600/-401	n.s. (p = 0.22)	n.s. (p = 0.34)	F _{1,29} = 5.04 p = 0.03 $\eta^2_g = 0.02$ (post-hoc t-tests ps > 0.11)
	mean	mean	
	difference = 2.04% CI 95% = [-1.3 5.37]	difference = -2.41% CI 95% = [-7.48 2.66]	
-400/-201	n.s. (p = 1)	n.s. (p = 1)	F _{1,29} = 8.37 p = 0.007 $\eta^2_g = 0.03$ (post-hoc t-tests ps > 0.31)
	mean	mean	
	difference = 1.7 ⁻¹⁰ % CI 95% = [-5.23 5.23]	difference = 1.7 ⁻¹⁰ % CI 95% = [-6.4 6.4]	
-200/-100	n.s. (p = 0.36)	F _{1,29} = 5.58 p = 0.03 $\eta^2_g = 0.03$ mean	n.s. (p = 1)
	mean	mean	
	difference = 2.22% CI 95% = [-2.73 7.17]	difference = -5.19% CI 95% = [-9.68 -0.7]	
0/200	n.s. (p = 0.71)	n.s. (p = 0.68)	n.s. (p = 0.93)
	mean	mean	
	difference = -0.93% CI 95% = [-6.04 4.19]	difference = -0.93% CI 95% = [-5.42 3.57]	
201/400	n.s. (p = 0.67)	n.s. (p = 0.39)	n.s. (p = 0.16)
	mean	mean	
	difference = 1.11% CI 95% = [-4.18 6.4]	difference = -2.22% CI 95% = [-7.41 2.96]	
401/600	n.s. (p = 0.18)	n.s. (p = 0.51)	n.s. (p = 0.31)
	mean	mean difference = 1.3%	
	difference = 3.89% CI 95% = [-1.88 9.66]	CI 95% = [-2.66 5.25]	
601/800	n.s. (p = 0.21)	n.s. (p = 0.59)	n.s. (p = 0.75)
	mean	mean	
	difference = 3.33% CI 95% = [-1.99 8.66]	difference = -1.48% CI 95% = [-7.04 4.07]	
801/1000	n.s. (p = 0.94)	n.s. (p = 0.46)	n.s. (p = 0.19)
	mean	mean	
	difference = 0.18% CI 95% = [-4.7 5.07]	difference = -2.41% CI 95% = [-8.98 4.17]	
1001/1200	n.s. (p = 0.28)	F _{1,29} = 7.9 p = 0.009 $\eta^2_g = 0.03$ mean	n.s. (p = 0.27)
	mean	mean	
	difference = 3.15% CI 95% = [-2.76 9.06]	difference = -6.86% CI 95% = [-11.84 -1.87]	
1201/1400	n.s. (p = 0.17)	n.s. (p = 0.55)	F _{1,29} = 4.79 p = 0.037 $\eta^2_g = 0.02$ (post-hoc t-tests ps > 0.12)
	mean	mean	
	difference = 3.52% CI 95% = [-1.63 8.67]	difference = 1.67% CI 95% = [-4 7.33]	
1401/1600	n.s. (p = 0.9)	F _{1,29} = 4.51 p = 0.04 $\eta^2_g = 0.03$ mean	n.s. (p = 1)
	mean	mean	
	difference = -0.37% CI 95% = [-6.56 5.82]	difference = -6.3% CI 95% = [-12.36 -0.24]	
1601/1800	n.s. (p = 0.07)	F _{1,29} = 10.1 p = 0.004 $\eta^2_g = 0.05$ mean	n.s. (p = 0.06)
	mean	mean	
	difference = 5.92% CI 95% = [-0.47 12.32]	difference = -8.15% CI 95% = [-13.4 -2.9]	
1801/2000	n.s. (p = 0.16)	F _{1,29} = 10.1 p = 0.001 $\eta^2_g = 0.07$ mean	n.s. (p = 0.43)
	mean	mean	
	difference = 3.14% CI 95% = [-1.3 7.6]	difference = -9.44% CI 95% = [-14.73 -4.16]	

percentage points).

4. Discussion

Our study was motivated by the recently observed unexpected SSVEP attenuation when multiple emotional as compared with neutral images were viewed in a rapid serial visual presentation at a 6 Hz rate (~167 ms per image), whereas a typical SSVEP enhancement was seen for affective relative to neutral images shown at a 4 Hz rate (250 ms per image). Here we investigated two critical open questions. First, we examined whether such a discrepancy may potentially be driven by a neutral image category more strongly capturing attention and, consequently, leading to larger SSVEP amplitudes with neutral picture content at a 6 Hz presentation rate. Secondly, the experiment aimed to test to what extent the time-scale of attentional deployment from the visual task, following the onset of distracting emotional cues, may be accompanied by a time-invariant reciprocal response enhancement in favor of emotionally arousing distracter-images. Combining our classical distraction paradigm with a rapid serial visual presentation (RSVP), the present design enabled an in-depth timepoint-by-timepoint analysis of the time-course of neural competition between the foreground detection task and rapidly presented emotionally arousing distracter-images, providing unique complementary information on their reciprocal temporal dynamics during sustained visual attention. In line with our hypotheses and the idea of limited capacity processing, we observed both prioritized processing of task-irrelevant emotional stimuli relative to neutral stimuli, irrespective of their presentation rate, followed by attentional deployment from the primary task in favor of emotional distracters.

The SSVEP response to the task-irrelevant RSVP background streams at 4 and 6 Hz started to significantly differ between neutral and unpleasant image streams around 225–270 ms after the onset of a change in emotional content. Replicating our recent 6 Hz RSVP results (Bekhtereva and Müller, 2017b; Bekhtereva et al., 2018), we found that changing a task-irrelevant 6 Hz background RSVP stream from neutral to unpleasant content elicited a significant reduction in SSVEP amplitude. Critically, this demonstrates that our effects were not attributable to the passive viewing of pictures per se, or random factors resulting from conditions of unconstrained attention, which may have potentially led to a stronger influence of more spontaneous emotional responses or situational factors (cf. Diers et al., 2014) in our earlier studies with 6 Hz presentation rate. Instead, the reverse emotional valence-driven SSVEP modulation was preserved even when participants were engaged in a demanding detection task while the 6 Hz RSVP in the background remained task-irrelevant. Thus, the reversal of the SSVEP amplitude modulation with emotional image content with 6 Hz RSVP remains a remarkable, consistent finding that could be replicated not only across different samples of subjects and experimental images, including positive emotional image valence (Bekhtereva et al., 2018) but also beyond a passive viewing paradigm, as the present experiment with a main primary visual task has shown. By contrast, yet also consistent with previous results from Bekhtereva et al. (2018), SSVEPs significantly increased in response to the change in content from neutral to negative in a 4 Hz RSVP at around 270 ms. These opposite patterns of SSVEP amplitude modulation prompt the inevitable question: do these presentation rates produce dissimilar effects on early visual cortex modulation with emotional content, altering which category – neutral or affective – receives preferential processing?

According to our hypothesis, if the reversed pattern of the SSVEP amplitude modulation for 6 Hz could be accounted for by neutral images being stronger competitors for processing resources, attracting more attention relative to emotional images, then there would be a greater visual distraction effect from the task with neutral rather than unpleasant distracters. Crucially, present findings show this was not the case. Regardless of whether distracting background RSVPs were presented at a

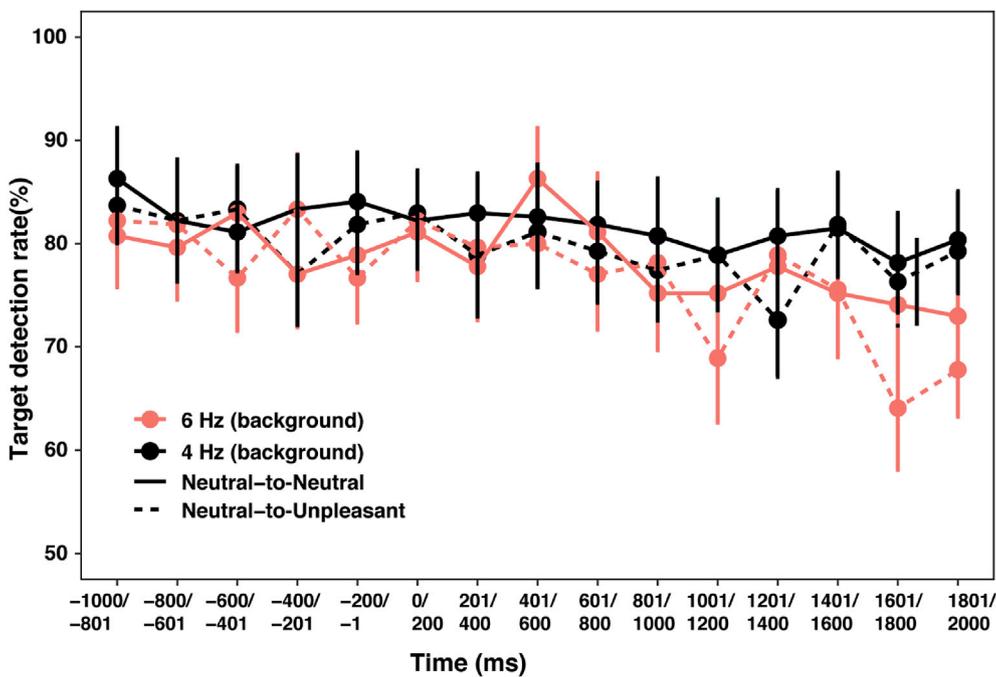


Fig. 6. Time-courses of target detection rates are shown in percentages for 15 consecutive time bins (five before and ten after the change in emotional content for background RSVPs), averaged for all participants. Hit rates are depicted for four experimental conditions: when neutral RSVP without a change in content (solid black line) or neutral RSVP with a change to unpleasant image content (dotted black line) were shown at 4 Hz in the background; when neutral RSVP without a change in content (solid red line) or neutral RSVP with a change to unpleasant image stream (dotted red line) were displayed at 6 Hz. Note that target detection rates were calculated as time-locked to the onset of the target events and, as a consequence, correspond to button presses that occurred about 600 ms later, on average. Error bars depict 95% confidence intervals and correspond to the within-subjects variability (Morey, 2008).

4 or 6 Hz rate, SSVEP amplitudes elicited by the foreground task at 15 Hz dropped significantly more when unpleasant as opposed to neutral distracters were displayed, resulting in a consistent emotion-driven attentional effect. As noted earlier, this reveals a more substantial deployment of attentional resources from the foreground task following a change from neutral to unpleasant background images. Hence, it unequivocally demonstrates that the direction of the SSVEP emotional modulation elicited by RSVP streams is irrelevant for biasing a neural competition for processing resources in favor of emotionally arousing content.

Importantly, we found that the onset of the shift of attentional resources from the foreground task toward RSVP streams of emotional distracters did not occur until ~ 470 ms following the onset of a change in emotional valence. Identically to our previous studies, the reduction in 15 Hz SSVEP amplitude lasted for several hundred milliseconds, subsiding around 1300 ms. It is remarkable that irrespective of whether a single (Bekhtereva and Müller, 2017a; Müller et al., 2011; Schönwald and Müller, 2014), or multiple images are presented as task-irrelevant, the observed attentional deployment effects evolve relatively late (~ 500 ms after picture/change onset), given that attentional capture by emotional stimuli is frequently shown to occur much faster (Carretie, 2014; Globisch et al., 1999; Pessoa, 2005).

We recently argued that this slow biasing of attention, as measured with SSVEPs, might represent reentrant feedback mechanisms from higher cortical areas involved in the decoding of emotional content (Bekhtereva et al., 2015; Bekhtereva and Müller, 2015; Keil et al., 2009). It is well established that affective content extraction for faces compared to IAPS scenes occurs at different speeds, as indexed by the modulations of the face-selective N170 component at ~ 170 ms following a face onset and the N1-EPN ERP complex developing around ~ 200 – 360 ms after a scene presentation (Aldunate et al., 2018; Olofsson et al., 2008; Schupp et al., 2012). Notably, we have recently demonstrated that a greater reduction of SSVEP amplitude elicited by the foreground task, signifying a larger withdrawal of attentional resources from the task toward a background emotional relative to a neutral image, occurred significantly later (i.e., by ~ 370 ms) when IAPS scenes compared to faces were presented in the background (Bekhtereva et al., 2015). In turn, this finding strongly suggested that temporal dynamics of an attentional shift toward affective images in early visual cortex relies on or follows the higher-order process of affective content identification. In corroboration

with that evidence, in the present study, the onset of the differential SSVEP modulation in response to unpleasant relative to neutral background RSVPs was observed at ~ 220 – 270 ms and was followed by the attentional resource shift as indexed by the modulation of the task-related 15 Hz SSVEP at about 470 ms, thus further strengthening the idea that SSVEP modulations with distracting emotional RSVPs are temporally linked to cortical emotional cue extraction as reflected in a typical N1-EPN complex latency (see above).

As discussed above, a biphasic process may explain the temporal pattern of our results. After the change to an emotional image category, a sensory gain mechanism first enhances the neural representation of emotional images in early visual areas, presumably via reentrant feedback projections from high-order cortical and subcortical networks involved in the processing of emotional image content (Amaral et al., 2003; Keil et al., 2009). The sensory gain in early visual areas via reverberating activity in a complex occipito-parieto-temporo-frontal network was recently demonstrated in fear learning using steady-state responses (Yuan et al., 2018). Following that initial sensory gain process, competitive interactions (Desimone and Duncan, 1995) between the spatially overlapping stimuli come into play triggering the shift of attentional resources away from the foreground task towards emotional background images. This explanation would be highly compatible with a framework in which additive influences and competitive interactions may occur between neural networks that guide top-down (voluntary) attention and motivated (emotional) attention in early visual processing areas (Pourtois et al., 2013).

Surprisingly, target detection accuracy did not reveal any reliable costs as a function of emotional valence of background images. Instead, an overall slight decrease in hit rates was observed in the late “post-change” time interval from ~ 1000 to ~ 1200 ms and from ~ 1400 to ~ 2000 ms for 6 Hz RSVPs relative to image streams presented at 4 Hz. One possibility is that it may perceptually have been slightly easier to distinguish the coherent target motion of flickering at 15 Hz RDKs when 4 Hz streams were shown in the background relative to the 6 Hz streams. However, despite being statistically significant, these differences in hit rates for 4 and 6 Hz in that time window were minimal and constituted only about 6–10 percentage points (Fig. 6). Moreover, in comparison with the consistently observed substantial drop by ~ 20 – 30% in hit rate percentage points, which was seen in our earlier studies shortly after the

onset of a concrete background image regardless of emotional valence (Bekhtereva and Müller, 2017a; Hindi Attar et al., 2010; Müller et al., 2011), the present differences seem negligible. Furthermore, a lack of target performance costs associated with a change from neutral to emotional RSVPs may likely be explained by the differences introduced in the current experimental design. Here, instead of displaying a single background distracter-image, we showed a continuous stream of multiple scenes that could rapidly change from neutral to a stream of emotional scenes. Therefore, it is likely that visual masking effects arising from showing a new image every presentation cycle (every ~167 or 250 ms) may have prevented image content from a more in-depth visual analysis that might be necessary to produce a substantial valence-dependent interference with target accuracy.

Our earlier research utilizing the same distraction paradigm could provide support for that line of interpretation. Unlike the current experiment, our previous studies displayed distracting emotional scenes for a much longer time. For example, a presentation of a phase-scrambled picture was followed by an onset of an intact neutral or emotional distracter-image that remained on the screen for ~2–3 s (Schönwald and Müller, 2014). Importantly, in that study, in addition to the deployment of attentional resources from the foreground task in favor of emotional distracters as indexed by the SSVEP modulation, we also observed robust amplitude modulations of the late positive potential (LPP). This ERP component with an enhanced positivity starting at around 500–800 ms following the onset of an intact emotional relative to neutral image serves as a neural signature of a more fine-detailed analysis of affective content during later stages of processing as well as sustained allocation of attention to emotional vs. neutral scenes (Hajcak et al., 2013; Liu et al., 2012). Likely, the LPP reflects a combination of top-down and bottom-up processes as demonstrated by stronger functional connectivity between prefrontal and occipital cortical areas (Moratti et al., 2011).

Consistent with this idea of a potentially weaker interference with behavioral accuracy due to a short picture exposure time in our current distraction paradigm are the findings from our recent experiment in which a background distracter-image appeared on the screen as briefly as ~130 ms, being subsequently masked (Bekhtereva and Müller, 2017a). Notably, in that experiment, no reliable LPP modulation was observed with background images, and, although target accuracy was briefly decreased between ~270 and 530 ms following the short presentation of an intact unpleasant relative to neutral picture, this drop in performance was more short-lived in comparison with the similar experiments when images stayed on the screen for a more extended period of time (Hindi Attar et al., 2010; Schönwald and Müller, 2014). Thus, it is conceivable that behavioral costs with distracting emotional scenes might be related to later, more elaborated processing stages of image content, whereas the perceptual shift in early visual processing resources as reflected in SSVEP modulations might be triggered by initial stages of emotional cue extraction (i.e., modulation of N1-EPN components).

Valence ratings would fit with the line of arguing above: in the post-experimental procedure, participants evaluated unpleasant images as slightly more negative when an image was displayed for 250 ms as compared with ~167 ms. As we have previously mentioned (Bekhtereva et al., 2018), presentation duration might influence the subjectively perceived emotional intensity of complex scenes (see also Codispoti et al., 2009). Nevertheless, the observed marginal absolute differences as 0.13 on average in valence ratings for unpleasant pictures (on a scale from 1 to 9) between 250 and ~167 ms exposure times cannot warrant a meaningful interpretation of this effect. More importantly, however, even with a picture presentation as short as ~167 and 250 ms followed by a mask, unpleasant scenes were evaluated as more negative than neutral visual scenes, thereby further supporting rapid affective content extraction with complex visual scenes.

5. Conclusions

The current study provides the first simultaneous timepoint-by-

timepoint assessment of the neural dynamics of sensory facilitation and attentional deployment in response to rapidly presented task-irrelevant emotional images during a sustained visual detection task. The results indicate two important points. First, our findings critically demonstrate that the SSVEP attenuation with affective images presented at 6 Hz is not a result of fundamentally distinct processing mechanisms rendering the neutral image category more superior in capturing attention and thus leading to the larger SSVEP amplitudes with neutral relative to emotional image content. Instead, it was affective distracter-images as opposed to their neutral counterparts that received preferential attentional processing and led to the greater deployment of visual processing resources from the foreground detection task, irrespective of whether they were presented at 6 or 4 Hz rate. This suggests that the directionality of valence-driven SSVEP amplitude modulation, that is, an increase or decrease in response to affective relative to neutral images may be less important than the presence of differential amplitude effects between the two valence categories per se. Second, we observed a biphasic process of first sustained sensory facilitation and a later occurring attentional deployment in favor of affective distracters, indicating that the relationship between sensory amplification and attentional bias with emotionally arousing scenes in early visual cortex is not a time-invariant reciprocal process. Thus, while detecting and processing emotional distracters can be fast, the resulting changes in attention are slower.

Conflicts of interest

The authors declare no conflict of interest.

Acknowledgments

The authors thank Renate Zahn for support in the data acquisition. The study was supported by Deutsche Forschungsgemeinschaft [MU972/22-2].

Appendix A

IAPS numbers of neutral pictures:

1122,1350,1645,1945,2036,2038,2191,2206,2273,2357,2377, 2445,2525,2745,5120,5130,5201,5500,5530,5535,7001,7010,7011, 7026,7030,7038,7041,7062,7081,7130,7136,7160,7161,7165,7300, 7491,7495,7512,7513,7546,7547,7560,7595,7632,8250,8325.

EmoPicS numbers of neutral pictures:

119,121,123,127,128,139,141,148,161,162,176,191,352,375.

IAPS numbers of unpleasant pictures:

1111,1113,1202,1220,2730,2981,3001,3019,3064,3103,3110, 3150,3190,3195,3212,3213,3250,3261,6021,6210,9002,9008,9031, 9075,9140,9163,9181,9342,9420,9471,9495,9570,9571,9590,9594, 9596,9600,9623,9635,9810,9901,9902,9911,9920,9930,9940.

EmoPicS numbers of unpleasant pictures:

216,232,233,234,235,236,240,241,243,248,321,325,326,327.

References

- Aldunate, N., López, V., Bosman, C.A., 2018. Early influence of affective context on emotion perception: EPN or early-N400? *Front. Neurosci.* 12, 708.
- Amaral, D.G., Behnia, H., Kelly, J.L., 2003. Topographic organization of projections from the amygdala to the visual cortex in the macaque monkey. *Neuroscience* 118 (4), 1099–1120.
- Andersen, S.K., Fuchs, S., Müller, M.M., 2010. Effects of feature-selective and spatial attention at different stages of visual processing. *J. Cogn. Neurosci.* 23 (1), 238–246.
- Baguley, T., 2012. Calculating and graphing within-subject confidence intervals for ANOVA. *Behav. Res. Methods* 44 (1), 158–175.
- Bakeman, R., 2005. Recommended effect size statistics for repeated measures designs. *Behav. Res. Methods* 37 (3), 379–384.
- Bekhtereva, V., Craddock, M., Müller, M.M., 2015. Attentional bias to affective faces and complex IAPS images in early visual cortex follows emotional cue extraction. *Neuroimage* 112, 254–266.
- Bekhtereva, V., Müller, M.M., 2015. Affective facilitation of early visual cortex during rapid picture presentation at 6 and 15 Hz. *Soc. Cogn. Affect. Neurosci.* 10 (12), 1623–1633.

- Bekhtereva, V., Müller, M.M., 2017a. Bringing color to emotion: the influence of color on attentional bias to briefly presented emotional images. *Cognit. Affect Behav. Neurosci.* 17 (5), 1028–1047.
- Bekhtereva, V., Müller, M.M., 2017b. Corrigendum to: affective facilitation of early visual cortex during rapid picture presentation at 6 and 15 Hz. *Soc. Cogn. Affect. Neurosci.* 12 (6), 1022–1023.
- Bekhtereva, V., Pritschmann, R., Keil, A., Müller, M.M., 2018. The neural signature of extracting emotional content from rapid visual streams at multiple presentation rates: a cross-laboratory study. *Psychophysiology* 55 (12), e13222.
- Benjamini, Y., Hochberg, Y., 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. Ser. B* 57 (1), 289–300.
- Bradley, M.M., Lang, P.J., 1994. Measuring emotion: the Self-Assessment Manikin and the semantic differential. *J. Behav. Ther. Exp. Psychiatry* 25 (1), 49–59.
- Capilla, A., Pazo-Alvarez, P., Darriba, A., Campo, P., Gross, J., 2011. Steady-state visual evoked potentials can be explained by temporal superposition of transient event-related responses. *PLoS One* 6 (1), e14543.
- Carrette, L., 2014. Exogenous (automatic) attention to emotional stimuli: a review. *Cognit. Affect Behav. Neurosci.* 14 (4), 1228–1258.
- Codispoti, M., Mazzetti, M., Bradley, M.M., 2009. Unmasking emotion: exposure duration and emotional engagement. *Psychophysiology* 46 (4), 731–738.
- 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 (1), 9–21.
- Delorme, A., Palmer, J., Onton, J., Oostenveld, R., Makeig, S., 2012. Independent EEG sources are dipolar. *PLoS One* 7 (2), e30135.
- Desimone, R., 1998. Visual attention mediated by biased competition in extrastriate visual cortex. *Philos. Trans. R. Soc. Lond. Ser. B Biol. Sci.* 353 (1373), 1245–1255.
- Desimone, R., Duncan, J., 1995. Neural mechanisms of selective visual attention. *Annu. Rev. Neurosci.* 18, 193–222.
- Deweese, M.M., Bradley, M.M., Lang, P.J., Andersen, S.K., Müller, M.M., Keil, A., 2014. Snake fearfulness is associated with sustained competitive biases to visual snake features: hypervigilance without avoidance. *Psychiatry Res.* 219 (2), 329–335.
- Deweese, M.M., Müller, M., Keil, A., 2016. Extent and time-course of competition in visual cortex between emotionally arousing distractors and a concurrent task. *Eur. J. Neurosci.* 43 (7), 961–970.
- Di Russo, F., Pitzalis, S., Aprile, T., Spitori, G., Patria, F., Stella, A., 2007. Spatiotemporal analysis of the cortical sources of the steady-state visual evoked potential. *Hum. Brain Mapp.* 28 (4), 323–334.
- Diers, K., Weber, F., Brocke, B., Strobel, A., Schönfeld, S., 2014. Instructions matter: a comparison of baseline conditions for cognitive emotion regulation paradigms. *Front. Psychol.* 5, 347.
- 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. *Behav. Res. Methods* 39, 175–191.
- Fuchs, S., Andersen, S.K., Gruber, T., Müller, M.M., 2008. Attentional bias of competitive interactions in neuronal networks of early visual processing in the human brain. *Neuroimage* 41 (3), 1086–1101.
- Globisch, J., Hamm, A.O., Esteves, F., Öhman, A., 1999. Fear appears fast: temporal course of startle reflex potentiation in animal fearful subjects. *Psychophysiology* 36 (1), 66–75.
- Gruss, L.F., Wieser, M.J., Schweinberger, S., Keil, A., 2012. Face-evoked steady-state visual potentials: effects of presentation rate and face inversion. *Front. Hum. Neurosci.* 6, 316.
- Hajcak, G., MacNamara, A., Foti, D., Ferri, J., Keil, A., 2013. The dynamic allocation of attention to emotion: simultaneous and independent evidence from the late positive potential and steady state visual evoked potentials. *Biol. Psychol.* 92 (3), 447–455.
- Hindi Attar, C., Andersen, S.K., Müller, M.M., 2010. Time course of affective bias in visual attention: convergent evidence from steady-state visual evoked potentials and behavioral data. *Neuroimage* 53 (4), 1326–1333.
- Jasper, H., 1958. The ten-twenty electrode system of the international federation. *Electroencephalogr. Clin. Neurophysiol.* 10, 371–375.
- Junghöfer, M., Elbert, T., Tucker, D.M., Rockstroh, B., 2000. Statistical control of artifacts in dense array EEG/MEG studies. *Psychophysiology* 37 (4), 523–532.
- Kaiser, J., Hertrich, I., Ackermann, H., Mathiak, K., Lutzenberger, W., 2005. Hearing lips: gamma-band activity during audiovisual speech perception. *Cerebr. Cortex* 15 (5), 646–653.
- 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 (5), 641–649.
- Keil, A., Costa, V., Smith, J.C., Sabatinelli, D., McGinnis, E.M., Bradley, M.M., 2012. Tagging cortical networks in emotion: a topographical analysis. *Hum. Brain Mapp.* 33 (12), 2920–2931.
- Keil, A., Gruber, T., Müller, M.M., Moratti, S., Stolarova, M., Bradley, M.M., 2003. Early modulation of visual perception by emotional arousal: evidence from steady-state visual evoked brain potentials. *Cognit. Affect Behav. Neurosci.* 3 (3), 195–206.
- Keil, A., Ihssen, N., 2004. Identification facilitation for emotionally arousing verbs during the attentional blink. *Emotion* 4 (1), 23–35.
- Keil, A., Sabatinelli, D., Ding, M., Lang, P.J., Ihssen, N., Heim, S., 2009. Re-entrant projections modulate visual cortex in affective perception: evidence from Granger causality analysis. *Hum. Brain Mapp.* 30 (2), 532–540.
- Lang, P., Bradley, M., Cuthbert, B.N., 2008. International Affective Picture System (IAPS): Affective Ratings of Pictures and Instruction Manual. University of Florida, Gainesville (FL). Technical Report A-8.
- Lawrence, M.A., 2013. *Ez: Easy Analysis and Visualization of Factorial Experiments*.
- Liu, Y., Huang, H., McGinnis-Deweese, M., Keil, A., Ding, M., 2012. Neural Substrate of the Late Positive Potential in Emotional Processing. *Journal of Neuroscience* 32 (42), 14563–14572.
- Moratti, S., Saugar, C., Strange, B.A., 2011. Prefrontal-occipitoparietal coupling underlies late latency human neuronal responses to emotion. *J. Neurosci.* 31 (47), 17278–17286.
- Morey, R.D., 2008. Confidence intervals from normalized data: a correction to Cousineau (2005). *Reason* 4 (2), 61–64.
- Morgan, S.T., 1996. Selective attention to stimulus location modulates the steady-state visual evoked potential. *Proceedings of the National Academy of Sciences of the United States of America* 93 (10), 4770–4774.
- Müller, M.M., Andersen, S., Trujillo, N.J., Valdes-Sosa, P., Malinowski, P., Hillyard, S.A., 2006. Feature-selective attention enhances color signals in early visual areas of the human brain. *Proceedings of the National Academy of Sciences of the United States of America* 103 (38), 14250–14254.
- Müller, M.M., Andersen, S.K., Hindi Attar, C., 2011. Attentional bias to briefly presented emotional distractors follows a slow time course in visual cortex. *J. Neurosci.* 31 (44), 15914–15918.
- Müller, M.M., Andersen, S.K., Keil, A., 2008. Time course of competition for visual processing resources between emotional pictures and foreground task. *Cerebr. Cortex* 18 (8), 1892–1899.
- Müller, M.M., Gundlach, C., 2017. Competition for attentional resources between low spatial frequency content of emotional images and a foreground task in early visual cortex. *Hum. Brain Mapp.* 38 (3), 429–443.
- Navarro, D.J., 2015. *Learning Statistics with R: A Tutorial for Psychology Students and Other Beginners, Version 0.5*. School of Psychology, University of Adelaide, Adelaide, Australia [Lecture notes].
- Norcia, A.M., Appelbaum, L.G., Ales, J.M., Cottareau, B.R., Rossion, B., 2015. The steady-state visual evoked potential in vision research: a review. *J. Vis.* 15 (6), 4.
- Olejnik, S., Algina, J., 2003. Generalized eta and omega squared statistics: measures of effect size for some common research designs. *Psychol. Methods* 8 (4), 434–447.
- Olofsson, J.K., Nordin, S., Sequeira, H., Polich, J., 2008. Affective picture processing: an integrative review of ERP findings. *Biol. Psychol.* 77 (3), 247–265.
- Pessoa, L., 2005. To what extent are emotional visual stimuli processed without attention and awareness? *Curr. Opin. Neurobiol.* 15 (2), 188–196.
- Peck, P., Schupp, H.T., Keil, A., Elbert, T., Junghöfer, M., 2009. Parallel processing of affective visual stimuli. *Psychophysiology* 46 (1), 200–208.
- 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 (3), 492–512.
- R Core Team, 2012. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna (Austria). Retrieved from: <http://www.R-project.org/>.
- Regan, D. (Ed.), 1989. *Human Brain Electrophysiology: Evoked Potentials and Evoked Magnetic Fields in Science and Medicine*. Elsevier, New York (NY).
- Schönwald, L.L., Müller, M.M., 2014. Slow biasing of processing resources in early visual cortex is preceded by emotional cue extraction in emotion-attention competition. *Hum. Brain Mapp.* 35 (4), 1477–1490.
- 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. *Biol. Psychol.* 91 (1), 81–87.
- Trauer, S.M., Andersen, S.K., Kotz, S.A., Müller, M.M., 2012. Capture of lexical but not visual resources by task-irrelevant emotional words: a combined ERP and steady-state visual evoked potential study. *Neuroimage* 60 (1), 130–138.
- Vialatte, F.-B., Maurice, M., Dauwels, J., Cichocki, A., 2010. Steady-state visually evoked potentials: focus on essential paradigms and future perspectives. *Prog. Neurobiol.* 90 (4), 418–438.
- Wessa, M., Kanske, P., Neumeister, P., Bode, K., Heissler, J., Schönfelder, S., 2010. EmoPics: subjective and psychophysiological evaluation of new imagery for clinical biopsychological research. *Zeitschrift für Klinische Psychologie und Psychotherapie, Supplementum* 1/11, 77 (available from: Michele Wessa, e-mail:wessa@uni-mainz.de).
- Wickham, H., 2009. *ggplot2: Elegant Graphics for Data Analysis*. Springer, New York.
- Wieser, M.J., Keil, A., 2011. Temporal trade-off effects in sustained attention: dynamics in visual cortex predict the target detection performance during distraction. *Cerebr. Cortex* 21 (12), 7784–7790.
- Wieser, M.J., Keil, A., 2014. Fearful faces heighten the cortical representation of contextual threat. *Neuroimage* 86, 317–325.
- Wieser, M.J., McTeague, L.M., Keil, A., 2012. Competition effects of threatening faces in social anxiety. *Emotion* 12 (5), 1050–1060.
- Wieser, M.J., Miskovic, V., Keil, A., 2016. Steady-state visual evoked potentials as a research tool in social affective neuroscience. *Psychophysiology* 53 (12), 1763–1775.
- 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 (3), 671–684.
- Yuan, M., Giménez-Fernández, T., Méndez-Bértolo, C., Moratti, S., 2018. Ultra-fast cortical gain adaptation in the human brain by trial-to-trial changes of associative strength in fear learning. *J. Neurosci.* 38 (38), 8262–8276.