Research paper

Acta Neurobiol Exp 2017, 77: 305–316

ACTA NEUROBIOLOGIÆ EXPERIMENTALIS

Top-down and bottom-up competition in visual stimuli processing

Tomasz S. Ligeza*, Agnieszka D. Tymorek, and Miroslaw Wyczesany

Psychophysiology Laboratory, Institute of Psychology, Jagiellonian University, Kraków, Poland,
** Email: tomasz.ligeza@uj.edu.pl*

Limited attention capacity results that not all the stimuli present in the visual field are equally processed. While processing of salient stimuli is automatically boosted by bottom-up attention, processing of task-relevant stimuli can be boosted volitionally by top-down attention. Usually, both top-down and bottom-up influences are present simultaneously, which creates a competition between these two types of attention. We examined this competition using both behavioral and electrophysiological measures. Participants responded to letters superimposed on background pictures. We assumed that responding to different conditions of the letter task engages top-down attention to different extent, whereas processing of background pictures of varying salience engages bottom-up attention to different extent. To check how manipulation of top-down attention influences bottom-up processing, we measured evoked response potentials (ERPs) in response to pictures (engaging mostly bottom-up attention) during three conditions of a letter task (different levels of top-down engagement). Conversely, to check how manipulation of bottom-up attention influences top-down processing, we measured ERP responses for letters (engaging mostly top-down attention) while manipulating the salience of background pictures (different levels of bottom-up engagement). The correctness and reaction times in response to letters were also analyzed. As expected, most of the ERPs and behavioral measures revealed a trade-off between both types of processing: a decrease of bottom-up processing was associated with an increase of top-down processing and, similarly, a decrease of top-down processing was associated with an increase in bottom-up processing. Results proved competition between the two types of attentions.

Key words: attention; top-down; bottom-up; competition; ERP

INTRODUCTION

Top-down and bottom-up competition

Attention is traditionally perceived as a filter which limits the overwhelming amount of information we receive every second. Stimuli perceptually 'compete' to be chosen by attention to get access to our cognitive resources. The choice of information for further processing is guided by: i) stimulus salience (related to emotional arousal, novelty, suddenness and general distinctiveness from other competing stimuli), and ii) behavioral relevance to the current task and established goals (Mayer et al. 2004). This dichotomy between automatically and voluntarily guided attention reflects the functional distinction between the bottom-up and top-down systems. Both of them constantly interact with each other in order to maintain the efficiency of goal-related actions whilst simultaneously monitoring the environment to allow for processing of unexpected but important events (Corbetta and Shulman 2002, Theeuwes 2010). This interaction can be seen as competition for limited perceptual resources, which can be drawn by either of the systems in certain conditions. Hence, the more resources utilized by the top-down system, the less are available for bottom-up processing (Berger et al. 2005, Hopfinger and West 2006, Okon-Singer et al. 2007). The aim of the present study was to check the competition between both systems on an electrophysiological level. By means of evoked potentials we measured how bottom-up and top-down processing of visual stimuli is affected by competition between systems during processing of visual stimuli.

Processing of visual stimuli depends on both the salience of the stimuli (bottom-up influence) and the volitional involvement of subjects (top-down influence). In laboratory settings, salience is most commonly manipulated by the valence of presented pictures, whose emotionally arousing material intensifies bottom-up processing (Dolan 2002). By contrast, volitional engagement is often studied with instructions that introduce task demands related to experimental stimuli. The task may be as simple as guiding the gaze to a particular part of the picture, or as complicated as performing mathe-

Received 6 April 2017, accepted 7 September 2017

Correspondence should be addressed to TS. Ligeza
Email: tomasz.ligeza@uj.edu.pl

matical operations that increase top-down processing during stimuli presentation. Although it has been consistently shown that salience affects the way stimuli are processed, there is mixed evidence regarding how such stimuli are processed while an additional task is imposed. Early findings suggest that salient stimuli have priority access to attention and that they can be processed without voluntary effort (Whalen et al. 1998). More recent evidence confirms that they can be processed non-volitionally, although they still require sufficient attentional engagement (Pessoa et al. 2005). For example, automatic processing of irrelevant but salient stimuli deteriorates when subjects are involved in very demanding tasks, but is preserved when tasks are relatively easy. Other studies have manipulated the load on working memory and, similarly, have shown that compared with low load conditions, a high load condition decreases processing of distracting stimuli (Lavie 2005, MacNamara et al. 2011, Van Dillen et al. 2009). This suggests that there is a competition between these two types of processing: the more top-down involvement, the weaker bottom-up driven processing. However, all of these previous studies have differentiated the level of engagement in a concurrent task in a binary manner: in low load conditions very easy competing tasks were used, while in high load conditions relatively difficult tasks were used. As a result, difficult tasks were confronted with easy tasks in which no deterioration of stimuli processing was observed. It is likely, however, that even easy tasks may affect the processing of task irrelevant stimuli to a lesser extent that is invisible at the behavioral level, but is still recognizable in the changes of electrophysiological activity. To observe these possible effects of concurrent tasks, we measured event related potentials (ERPs) in response to two types of stimuli while manipulating engagement of the attentional networks in a specially designed procedure.

A more ecologically valid task designed to study competition

Specifically, we designed a visual attention-distraction task (VDT) (Kossowska et al. 2015) in which subjects freely observed a screen to identify a small, supraliminally presented letter against standardized emotional pictures. These background pictures belonged to one of three conditions (PICTURE factor): negative (NEG, salient stimuli with a high emotional value), neutral (NEU, low motivational value), and scrambled images (SCR, pictures with no meaningful content and a low motivational value). The salience of the pictures was intended to influence bottom-up attention. On the other hand, three task conditions were introduced to influence top-down processing (CONDITION factor). They were signalized by a specific visual cue before each trial: i) search for and identify the letter in an expected location, determined by an arrow (DT), ii) identify the letter in an unknown location (UN), or iii) ignore the letter and do not respond (NR). Subjects were asked to refrain from responding in case they had not spotted the letter in a particular trial where a response was required.

This design allowed us to provide participants with more ecologically valid conditions than typical experimental settings. In order to distinguish between the central field and peripheral effects, a great deal of research uses paradigms which differ from natural settings and require participants to keep their eyes fixed on a particular point. In our daily lives we are not restricted to a specific part of the visual field and we are able to scan the entire visual field by performing saccades freely throughout the whole visual field. Importantly, in our design we had two types of visual stimuli: pictures and letters, each of which served to manipulate another attentional network. By manipulating the valence of the pictures (NEU, SCR, NEG) we impacted the bottom-up system activation, whereas by manipulating task conditions (DT, UN, NR) we impacted the activity of the top-down system. Both top-down and bottom-up processes are related to each other and compete for the same limited resources. Therefore, impacting the top-down attention has the opposite effects on the bottom-up attention and vice versa. As such, manipulating the letter task and the background of presented pictures allowed us to investigate the competition between bottom-up and top-down attention.

Regarding picture conditions, we assumed the highest involvement of the bottom up attention during presentation of negative pictures (NEG) and the lowest during presentation of neutral pictures (NEU). Scrambled pictures (SCR) were introduced here as non-emotional, meaningless perceptual content that was expected to evoke weaker bottom-up attention involvement than the effect of meaningful (NEU) or emotionally eliciting (NEG) stimuli. In regard to the task conditions, identification of a letter in an expected location (DT) was assumed to narrow the scope of attention by engaging the top-down (selective) attentional system to the highest degree (focusing attention on part of a visual scene, towards the expected letter). The condition in which the location of the letter (UN) was unknown also demanded activation of the top-down system (searching the visual field), but it simultaneously should made subjects more prone to distraction by task-irrelevant stimuli (background pictures). NR, the last condition (with the letter being task-irrelevant), was assumed not to involve top-down attention to a significant extent and, as a result, minimize processing of the letter (control condition for our top-down manipulation) and maximize processing of the picture (reflecting bottom-up system activation). As these conditions differ in the intensity of competition between top-down and bottom-up attention, our way to discover the effects of this competition on processing of visual stimuli was to measure ERP response in a specific task design. Especially, we were interested in measuring the response to letters while manipulating the valence of background pictures and measuring ERP components in response to pictures while manipulating the conditions.

Typical ERP components in studying of attention

Studies of attention typically examine several ERP components to learn about different stages of information processing, with P1, N1, P3 and LPP being the most important. P1 amplitude (80-140 ms over the posterior sites) increases when attention is allocated to the stimulus location, marking automatic suppression of stimuli outside the focus (Dennis and Chen 2007, Grzybowski et al. 2014). N1 (170-210 ms, parieto-occipital) reflects early processes associated with higher-level discrimination and the depth of early attentional capture. It also reflects correctly located attention towards upcoming, task-relevant stimuli (Vogel and Luck 2000). These early components are also often reported to reflect the biological significance of stimuli, which results in their deeper processing (Hart et al. 2012, Olofsson et al. 2008) (Hart et al. 2012, Olofsson et al. 2008). P3 is a heterogeneous component (250-500 ms, midline fronto-central or centro-parietal) that is thought to reflect the informational content of stimulus, memory operations, and decisive processes of selection and preparation of the correct behavioral response. It is also modulated by frequency of stimuli presentation and probability of its appearance (Gmaj et al. 2016, Verleger et al. 2005, Wronka et al. 2012). Additionally, the level of actual processing depth for motivationally-arousing stimuli (emotional slides in our case) can be inferred from the late positive potential (LPP). This centro-parietal component starts approximately 600ms after stimulus onset, lasts up to a few seconds and reflects elevated processing of stimuli (Hajcak et al. 2009, Jaśkiewicz et al. 2016).

Hypothesis

We assumed that our manipulations would reveal the trade-off between both networks: the greater involvement of the top-down system and the lesser activation of the bottom-up system (and vice versa). Thus, our aim was not to replicate the well-known ERP effects of valence and the effects of attended/unattended stimuli, but to check how increased recruitment of one system affects the effectiveness of the other: how a background picture influences the processing of letters and how attending letters influences the processing of background pictures.

To realize this aim we conducted two analyses. Firstly, we measured ERPs in response to picture presentation (bottom-up influences) while manipulating top-down attention (CONDITION). Secondly, we measured ERPs in response to letters (top-down influences) while manipulating bottom-up attention (PICTURE). We expected that attention-related EPRs in response to picture presentation (bottom-up influences) would be the most pronounced in the NR condition (decreased activation of top-down system), lower in the UN condition (moderate activation of top-down system), and lowest in the DT condition (increased top-down activation of top-down system). On the other hand, we expected that ERPs in response to letter presentation (top-down influences) would be the most pronounced in the SCR condition (decreased activation of bottom-up system), lower in the NEU condition (moderate activation of bottom-up system), and lowest in the NEG condition (increased bottom-up activation of bottom-up system).

METHODS

Participants

52 volunteers (41 women, age mean/SD: 23.0±2.5 yrs) participated in the study. All were medication-free with no reported history of any neurological or psychiatric disorders or substance abuse, and had normal or corrected-to-normal vision.

Apparatus and materials

A total of 180 pictures were used in the study. 120 pictures were selected from the Nencki Affective Picture System (NAPS) (Marchewka et al. 2014); of these, 60 depicted neutral scenes (NEU, household objects, neutral landscapes, people in everyday activities or ordinary relationships, neutral animals) and 60 depicted unpleasant scenes (NEG, violent images, sad people, animal mutilation, surgical procedures, accidents). The two categories differed on normative ratings of valence (mean/SD: 6.35±0.87 for neutral; 3.38±1.01 for negative pictures) and arousal (4.47±0.66 for neutral; 6.2±0.61 for negative pictures). An additional 60 pictures with potentially neutral valence ratings were scrambled (SCR) by randomly rearranging 0.5 cm square fragments of each original picture. All pictures were normalized in terms of their averaged brightness using ImageMagick processing software with histogram matching method.

The average brightness histogram was calculated from five randomly selected pictures. Then, the individual histograms of all pictures were matched with the average one (using joint rgb pixels intensity). As a result, all pictures used in the procedure, regardless of their category (NEU, NEG, SCR), had the same average brightness. The task was presented on a computer with a 61 cm LCD monitor. Each photo was presented full-screen on the monitor at a viewing distance of approximately 60cm and 50° of the horizontal visual angle.

Procedure

The procedure was compliant with the directives of the Helsinki Declaration and approved by the Ethical Committee of the Institute of Psychology, Jagiellonian University. Upon arrival the participants were provided with a brief description of the experiment (registering brain activity while viewing pictures and performing simple cognitive tasks). They were asked to complete an informed consent form. Finally, the EEG equipment was mounted and the participants received detailed task instructions. They were told that they would see cues (arrow, question mark or cross) followed by pictures and that the cues determined how they should react to the upcoming picture. Depending on the cues, the task was to: i) search for and identify the letter superimposed on the picture in an expected location determined by an arrow (arrow cue, DT condition); ii) identify the letter in an unknown location (question mark cue, UN condition); or iii) look at the picture but ignore the letter and respond to neither (cross cue, NR condition). To identify the letter, participants had to press the appropriate character on the keyboard (either "z" or "m") as soon as possible. The location of the letter stimuli was randomized in each trial, ranging from 20 to 95% of the screen dimension measured from its center to the edge, in both horizontal and vertical directions. There were no specific instructions regarding fixation point, as we wanted to make our procedure more natural. We intended the images to be freely and spontaneously examined. After receiving the instructions, the participants underwent a short training session. In the experimental session all participants attended 180 trials with breaks after every 36 trials (5 blocks). The timing of the presentation was as follows: cue 0.5s, blank screen 1.5 s, picture 3.8–4.8 s. Additionally, after 1.2–2.2 s after the picture onset, a small supraliminal letter was presented on the picture at either a random location (UN, NR condition) or at a location determined by the cue (DT condition). Each letter was shown for 600ms. After the offset of the letter, the picture was presented for further 2 s. The screen background was grey during the procedure. Mean luminance

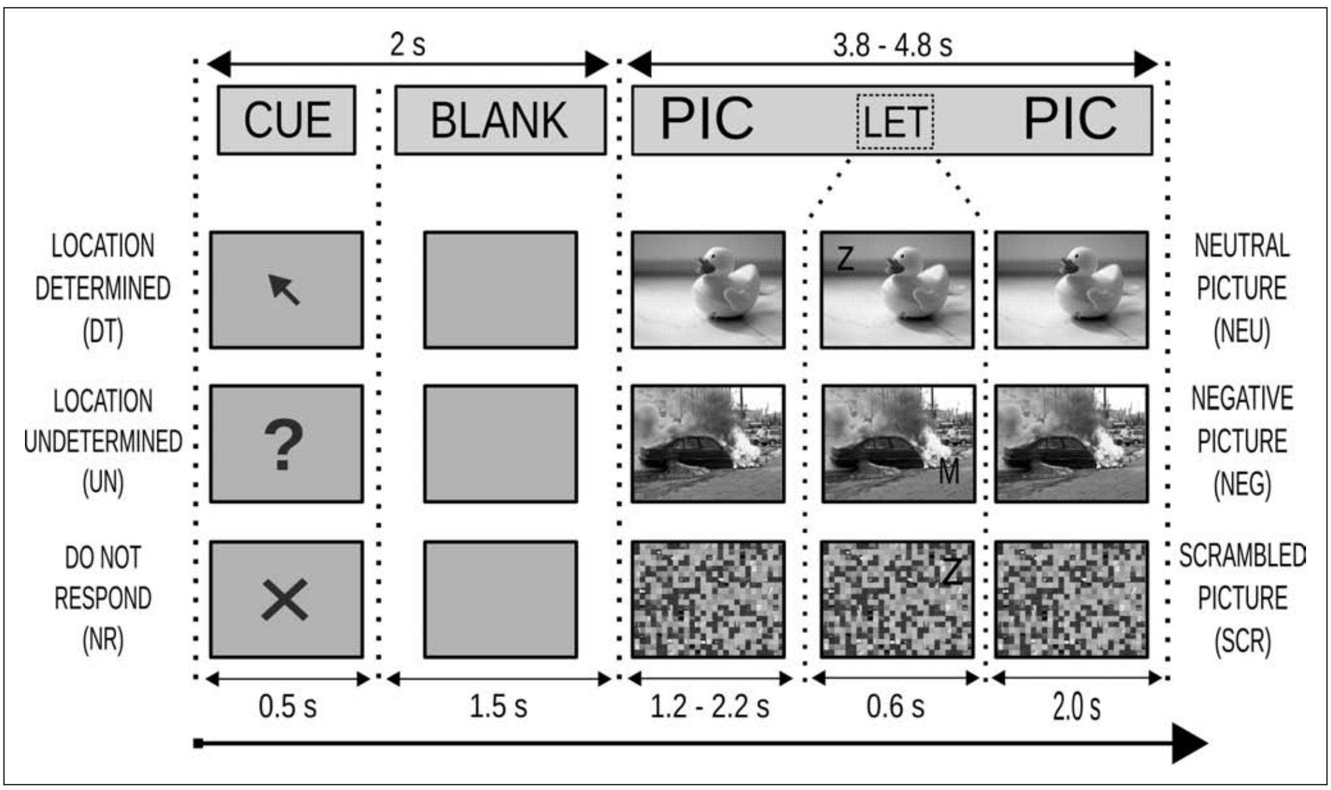

Fig. 1. The timeline of the experimental procedure with a scheme of three example trials.

for the blank screen was 16.10 cd/m^2, whereas mean luminance for pictures was 20.13 cd/m^2. The timeline of the procedure is presented in Fig. 1.

Behavioral measures

Two behavioral parameters were analyzed: reaction time (RT) and ratio of error to total responses (ER). Statistical analyzes were performed using SPSS Statistics version 21 (IBM, Armonk, NY, USA). The RT data was log-transformed because data distribution was skewed. A (3) CONDTION (DT *vs.* UN *vs.* NR) × (3) PICTURES (NEU *vs.* SCR *vs.* NEG) repeated measures ANOVA was then conducted, followed by Bonferroni corrected post-hoc tests. Due to the ER score distribution, we used the non-parametric Friedman's test.

Psychophysiological recordings

The EEG recording was carried out using the Biosemi ActiveTwo device with 64 electrodes placed on a head cap and two more electrodes for linked mastoid reference. An additional four leads were used to record the signal generated by blinks and eye movements. The hardware low-pass filtering was set to 410 Hz with 5th order sinc response filter. Following downsampling the final sampling rate of the recorded signal was 256 Hz. Preprocessing and artifact rejection was performed using EEGlab version 13, an interactive Matlab toolbox (Delorme and Makeig 2004). The signal was off-line filtered using 0.1 high-pass and 46 Hz low-pass filters (zero-phase Hamming windowed sinc finite impulse response digital filters, order 8448 and 74, respectively). Artifact rejection was carried out using Independent Component Analysis decomposition to identify and subtract eye-blinks and eye movement sources. The signal was segmented into 0-1000 ms time epochs relative to stimulus onset with baseline correction from -100 ms. Based on prior studies and visual inspection of the grand averages, we selected electrodes related to ERPs of our interest: PO7, PO8 for the N1 and the P1 components (e.g. Clementz et all. 2004); Fz, FCz for the P3a component; Cpz, Pz for the P3b component (e.g. Johnson 1993, Wronka et al. 2012) and Cz, CPz, Pz for the LPP component (e.g. Wyczesany and Ligeza 2017). Time boundaries were defined according to visual inspection of grand average waveforms. These time windows were centered around the grand-averaged peaks of N1, P1, P3a, P3b components on previously selected electrodes. LPP was defined as sustained positivity that occurs after P3b component. Finally, averaged amplitudes of ERP components were extracted for selected time-windows and electrodes separately for pictures and letters. Pictures: N1 (150–190 ms; PO7, PO8); P3 (310-400 ms; Fz, Fcz); LPP (600–1000 ms; Cz,Cpz, Pz). Letters: P1 (120–150 ms; PO7, PO8), N1 (175–210 ms; PO7, PO8), P3 (330–415 ms; Cpz, Pz). P1 was analyzed for letters only, as early attention allocation was not considered for pictures, while LPP was analyzed only for pictures as a marker of emotional processing (this component is not analyzed for letters). As responses to pictures were automatic, the P3 component was visible earlier and was located more anteriorly (P3a). Conversely, as letters required a response, the component was noticeable later and was located more posteriorly (P3b).

EEG data

Statistical analyzes were performed using SPSS Statistics version 21 (IBM, Armonk, NY, USA). A series of (3) CONDITION (DT *vs.* UN *vs.* NR) × (3) PICTURE (NEU *vs.* SCR *vs.* NEG) X ELECTRODE (all selected electrodes for a given ERP) repeated measures ANOVA were carried out. The ANOVAs were followed by post-hoc tests. Where applicable, the Greenhouse-Geisser correction was used. All the comparison p-levels reported below were subject to Bonferroni correction. The results that directly reflect our hypotheses (effects of letters presentation on processing of pictures and effects of picture presentation on processing of letters) are presented first. Then, the other relevant effects are presented. Results concerning our hypotheses are presented in Table I and ERP waveforms are presented in Fig. 2.

Table I. ERP Amplitude Data (Mean±SD)

	Pictures				Letters		
Component	DT	UN	NR	Component	NEG	NEU	SCR
N1	4.53±0.61	5.10±0.70	5.37±0.59	P1	0.10±0.17	0.80±0.17	0.14±0.18
P3a	-0.44±0.83	-6.55±0.92	-5.41±0.88	P3b	9.87±0.68	11.97±0.75	11.46±0.65
LPP	-2.78±0.59	-2.20±0.68	0.82±0.59				

RESULTS

Behavioral data

The main effect of CONDITION was revealed ($F_{1,51}$=98.77, P<0.001, η^2=0.66) with significantly (P<0.001) shorter reaction times for DT (mean/SE: 0.596±0.16 s) than the UN condition (0.662±0.16 s), and the main effect of PICTURE ($F_{2,50}$=7.90, p=0.001, η^2=0.24) with significant differences between NEU (0.618±0.16 s) and NEG (0.632±0.17 s; p=0.007) as well as between NEU and SCR (0.634±0.16 s; P=0.005). Interaction PICTURE × CONDITION was non-significant. For error rates there was no statistically significant difference between picture types (χ^2(2)=2.61; P=0.27). However, there was a substantial difference between CONDITION conditions (Z=-4.70; P<0.001) with a higher number of errors for the UN (6.52) than the DT condition (1.81).

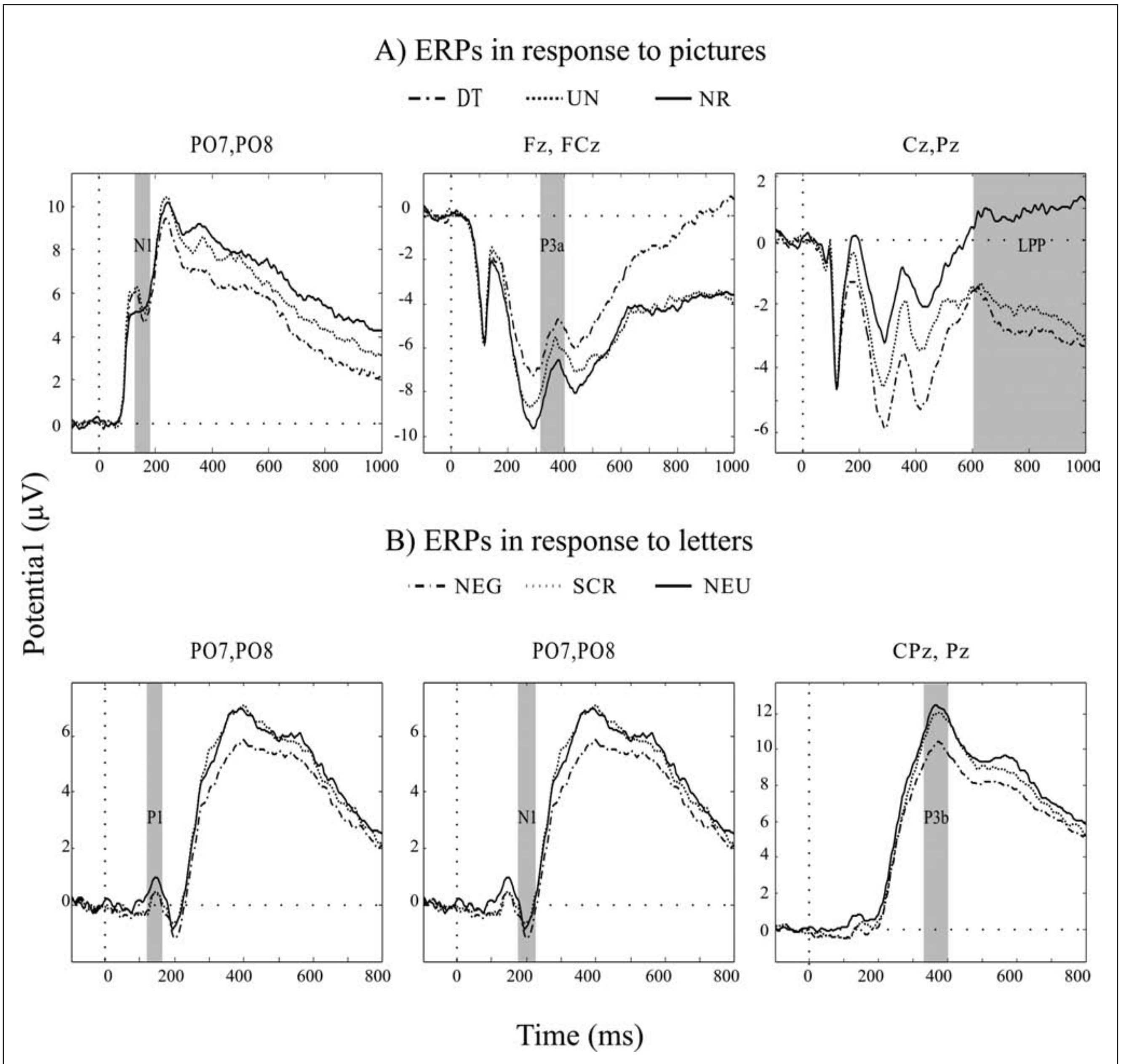

Fig. 2. ERP waveforms. The figure shows stimulus-locked grand average ERP waveforms in response to: (A) pictures (N1, P3a and LPP) during three task condition: DT- identify the letter in an expected location determined by an arrow, UN- identify the letter in an unknown location, NR - ignore the letter and do not respond; (B) in response to letters (P1, N1, P3b) presented against different backgrounds: NEG- negative picture, SCR- scrambled picture, NEU- neutral pictures. Shadings highlight the components time windows.

EEG data

Effect of condition on evoked responses to pictures

N1 (150–190 ms)

The main effect of CONDITION was found ($F_{2,49}$=7.43, P=0.002, η^2=0.23). The highest N1 amplitude was found for the DT condition (4.53±0.61 μV) and was significantly different from both NR (5.37±0.59 μV; p=0.002) and UN (5.10±0.70 μV; P=0.043).

P3 (310–400 ms)

The main effect of CONDITION was found ($F_{2,50}$=9.80, P<0.001, η^2=0.28) with a significantly higher amplitude in the NR condition (-5.41±0.88 μV) than both the UN (-6.55±0.92 μV; P=0.05) and the DT (-7,44±0.83 μV; P<0.001) conditions. The difference between the UN and DT conditions was not significant (P=0.079).

LPP (600–1000 ms)

LPP showed the main effect of CONDITION ($F_{2,50}$=34.58, P<0.001, η^2=0.58), with the highest amplitude in the NR condition (0.82±0.59 μV) in comparison to both the UN (-2.20±0.68 μV; P<0.001) and the DT conditions (-2.78±0.59 μV; P<0.001).

Effect of valence on evoked responses to letters

P1 (120–150ms)

The main effect of PICTURE ($F_{2,50}$=10.3, P<0.001, η^2=0.29) was found with the highest amplitude for NEU (0.80±0.17 μV) compared to both NEG (0.10±0.17 μV; P<0.001) and SCR (0.14±0.18 μV; P=0.007).

N1 (175–210 ms)

No main effect of PICTURE was observed.

P3 (330–415 ms)

The main effect of PICTURE was found ($F_{2,34}$=14.60, P<0.001, η^2=0.46) with a significantly lower amplitude for NEG pictures (9.87±0.68 μV) in comparison to both NEU (11.97±0.75 μV; P<0.001) and SCR (11.46±0.65 μV; P=0.002) pictures. Difference between NEU and SCR was non-significant.

Other effects related to evoked responses to pictures

N1 (150–190 ms)

Main effect of PICTURE was found ($F_{2,49}$=26.26; P<0.001, η^2=0.52). The highest N1 amplitude was found for the SCR pictures (3.38±0.63 μV) and was significantly different from both NEG (5.93±0.66 μV; P<0.001) and NEU (5.69±0.65 μV; P<0.001) pictures. The difference between NEG and NEU was non-significant.

PICTURE × CONDITION interaction was significant ($F_{4,47}$=4.54; P=0.003, η^2=0.28). Post hoc analysis revealed that influence of CONDITION was significant only for SCR pictures with significantly higher N1 amplitude for the DT (2.46±0.66 μV) in comparison to the NR (4.32±0.62 μV; P<0.001) and for UN (3.35±0.73 μV; P=0.05). The difference between NR and UN was non-significant.

ELECTRODE × CONDITION interaction was significant ($F_{2,49}$=3.46; P=0.039, η^2=0.124). For PO7 electrode higher N1 amplitude was found for DT (4.62±0.61 μV) in comparison to UN condition (5.31±0.68 μV; P=0.039). There were no differences between either NR (5.17±0.52 μV) and UN or NR and DT. For PO8 electrode we found a significantly higher N1 amplitude for DT (4.44±0.74 μV) in comparison to NR (5.56±0.75 μV; P<0.001). Other effects were non-significant.

P3 (310–400 ms)

Main effect of PICTURE was found ($F_{2,50}$=27.50; P<0.001, η^2=0.52).

The highest P3 amplitude was found for the SCR pictures (-4.35±0.82 μV) and was significantly different from both NEG (-7.56±0.98 μV; P<0.001) and NEU (-7.49±0.84 μV; P<0.001). Difference between NEG and NEU was non-significant.

PICTURE × CONDITION interaction was non-significant. ELECTRODE × CONDITION interaction was significant ($F_{2,50}$=6.56; P=0.003, η^2=0.21). For Fz electrode higher P3 amplitude was found for NR (-5.57±0.90 μV) in comparison to DT (-7.34±0.84 μV; P=0.001). There were no differences between either UN (-6.53±0.95 μV) and NR or UN and NR. For FCz electrode the highest P3 amplitude was for NR (-5.26±0.88 μV, P=0.001) in comparison to both UN (-6.57±0.90 μV; P=0.024) and DT(-7.54±0.83 μV; P<0.001). Difference between UN and DT conditions was also significant (P=0.044).

LPP (600–1000 ms)

Main effect of PICTURE was found ($F_{2,50}$=24.44, P<0.001, η^2=0.49). The highest LPP amplitude was found for the NEG pictures (0.31±0.65 μV) and was significantly different from both NEU (-1.97±0.59 μV; P<0.001) and SCR (-2.50±0.58 μV; P<0.001). There was no significant difference between NEU and SCR.

PICTURE × CONDITION interaction was non-significant.

ELECTRODE × CONDITION interaction was non-significant.

Other effects related to evoked responses to letters

P1 (120-150 ms)

Main effect of CONDITION was found ($F_{2,50}$=4.58, P=0.015, η^2=0.16). Post hoc revealed statistically significant difference of the P1 amplitude between the NR

(0.01±0.144 μV) and the UN condition (0.57±0.19 μV; P=0.013). The DT condition (0.45±0.18 μV) was not significantly different from either the NR or the UN condition.

PICTURE × CONDITION interaction was non-significant.

ELECTRODE × PICTURE interaction was non-significant.

N1 (175–210 ms)

Main effect of CONDITION was found ($F_{2,49}$=8.33, P=0.001, η^2=0.25). The highest N1 amplitude was found for the DT condition (-1.08±0.34 μV) and it was significantly different from both the UN (-0.48±0.37 μV; P=0.030) and the NR(0.06±0.19 μV; P=0.001). There was no significant difference between the UN and the NR conditions.

PICTURE × CONDITION interaction was non-significant.

ELECTRODE × PICTURE interaction was non-significant.

P3 (330-415 ms)

Main effect of CONDITION was found ($F_{2,34}$=71.28, P<0.001, η^2=0.81). The highest P3 amplitude was found for the DT condition (15.49±0.97 μV) and was significantly different from both the UN (13.31±0.86 μV; P=0.003) and the NR (4.49±0.46 μV; P<0.001). There was also significant difference between the UN and the NR conditions (P<0.001). PICTURE × CONDITION interaction was significant ($F_{2,32}$=3.41, P<0.020, η^2=0.30). For each condition pattern of differences among pictures was different. For the DT condition P3 amplitude was significantly higher for NEU pictures (16.54±1.10 μV) than for NEG pictures (14.78±0.97 μV; P=0.006) and SCR pictures (15.95±1.04 μV; P=0.05). For the UN condition the highest P3 amplitude was found for NEU pictures (15.32±1.03 μV) and was significantly different from both NEG (11.99±0.86 μV; P<0.001) and SCR (12.64±0.95 μV; P=0.004). There was no significant difference between NEG and SCR pictures. For the NR condition the lower P3 amplitude was found for NEG (3.44±0.57 μV) than for SCR pictures (5.97±0.80 μV; P=0.003). NEU pictures (5.06±0.54 μV) was not significantly different from either NEG or SCR pictures. ELECTRODE × PICTURE interaction was non-significant.

DISCUSSION

The study investigated the competition between top-down and bottom-up attention using evoked potentials in a more ecologically valid procedure than previous research. We assumed here that manipulations of letter task conditions (identifying letter in known/unknown position / not responding to letters) changed the degree to which top-down processing was activated. Similarly, it was assumed that manipulation of salience of background pictures (scrambled/neutral/negative) affected engagement of bottom-up processing. Thus, to check how intensification of top-down attention influences bottom-up processing, we measured ERP responses for pictures while manipulating task conditions, and, conversely, to check how intensification of bottom-up attention influences top-down processing, we measured ERP responses for task-letters while manipulating salience of pictures presented as a background. As expected, most of the ERP measures (P1, P3 and LPP components) revealed a decrease of bottom-up attention associated with an increase of top-down processing and, similarly, a decrease of top-down attention associated with an increase in bottom-up processing. The results concerning the N1 component were either contrary to our expectation (N1 in response to pictures) or not significant (N1 in response to letters). Overall, these mutually-inverted patterns of P1, P3 and LPP components proved the constant competition between the two types of attentions in a relatively easy task.

ERPs in response to pictures

The hypothesis regarding processing the pictures with respect to the additional letter task was confirmed: additional engagement of top-down attention decreased bottom-up influences. This was revealed by both the P3 and LPP components in response to pictures. The mean amplitude of these components decreased when subjects had to respond to the letters (either in predefined, DT or undetermined conditions, UN) compared to the condition in which they had to ignore the letters. Still, the exact patterns of the results were not entirely as expected. In line with the hypothesis, the P3 component in automatic response to pictures in the DT condition (maximal top-down letter-task influence) decreased compared to the NR condition (minimal top-down influence). However, the response to pictures in the UN condition was significantly different neither from the DT nor the NR conditions (the mean amplitude of the component was between the two conditions). It is possible that the UN task, as intended, allowed for similar intensification of both top-down and bottom-up processing; however, the difference between the rest conditions was too small. It is also possible that the UN condition engaged top-down attention to a similar extent as the DT condition. Although searching an image for a letter in an unexpected location could increase bottom-up attention more than in the DT condition, at the same time, the process of searching could be more challenging for top-down attention than guiding attention in a predefined manner.

As a result, both types of attention could be engaged similarly. This latter interpretation is confirmed by a pattern of results observed for the LPP component, in which the responses in both the DT and UN conditions were lower than in the NR condition. This suggests that both conditions of the letter tasks engaged top-down processing to a similar extent. Aside from these minor ambiguities regarding the two conditions representing the letter task, the prevailing pattern of P3 and LPP components in response to pictures confirms the expected dependency: during involvement of top-down processing (DT or UN) the decrease of bottom-up response to pictures was clearly visible.

Contrary to our expectations, the N1 component was most pronounced in the DT condition. This might suggest that the increased N1 in response to pictures reflected not the bottom-up process related to actual processing of the whole picture, but another cognitive process. As attention in the DT condition was already narrowed to a selected part of the visual field, the component could represent an automatic, bottom-up process of orienting attention towards the whole picture. In this vein, the N1 component could mark the bottom-up process related to automatic orienting of attention towards appearing stimuli. Such a role of the N1 component is often reported in the literature (Luck et al. 1990, Luck et al. 1994, Rugg et al. 1987). Alternatively but not contradictory to this reasoning, narrowing attention to a particular field could per se increase the N1 component. One study manipulated global *vs.* local attention to pictures and reported increased N1 component when attention was local (Gable and Harmon-Jones 2012). The latter interpretation is further strengthened by a significant PICTURE × CONDITION interaction. Pattern of results observed for the N1 component was significant only for scrambled pictures, whereas differences between conditions were not significant for other types of pictures. As it is harder to distinguish a letter from the scrambled than non-scrambled background, the N1 could reflect a more narrowed attention to the part of the picture where a letter was supposed to appear. Less demanding backgrounds (negative, neutral) did not demand such narrowed attention. Taken together, the results of P3 and LPP but not the N1 component were in line with our hypothesis regarding the competition between bottom-up and top-down attention.

ERPs in response to letters

The hypothesis regarding the influence of the type of background picture on processing letters was also confirmed: salient background pictures captured bottom-up attention at the expense of letter processing. Thus, compared with the case described in the previous paragraph in which ERPs in response to pictures were decreased during intensification of top-down processing, here we observed a reversed pattern in which intensifying of bottom-up influences (represented by the valence of the picture) decreased top-down attention focused on the letters. Also in this case, these expected patterns of results were observed for most, but not all, of the ERP components. As expected, both P1 and P3 components in response to letters were maximal while the background picture of the letter was neutral compared to the negative and scrambled ones. Significant interaction PICTURE × CONDITION for P3 component in response to letters suggests that above mentioned pattern of the results applied to DT and UN conditions, whereas for the NR condition this pattern was slightly different. For the NR condition, P3 in response to letters was greater for scrambled relative to negative background with others differences being non-significant. It suggests that when participants did not respond to letters, the letters attracted the attention to a greater extend when the background was meaningless (scrambled).

Unexpectedly, the N1 component in response to the letters was not significantly different for different pictures types. Thus, similarly to the analysis of ERPs in response to pictures, the N1 component did not reveal signs of the trade-off between top-down and bottom-up processing, that was observed for other components (P1, P3 and LPP). While the effects of the N1 component in response to pictures was contrary to our expectations, in case of responses to letters these effects were not significant. Overall, this suggests that the N1 is not a suitable component to study the competition between top-down and bottom-up processing.

Interestingly, the effects of P1 and P3 components in response to letters during presentation of varied pictures differed slightly, showing the diversified impact of scrambled and negative images on processing letters. The magnitude of the P1 component in response to letters was identical for negative and scrambled backgrounds, but lower than for neutral pictures. This suggests that at the early processing stage (marked by the P1 component), negative and scrambled pictures capture attention to a similar extent. On the other hand, the P3 component was the same for neutral and scrambled backgrounds as for negative pictures. This, in turn, suggests that during the subsequent phase of picture processing, increased processing of negative pictures is maintained, whereas processing of scrambled pictures drops. Such a discrepancy might be explained by differences in the complexion and content displayed by the scrambled and negative pictures. The scrambled images are comprised of a colorful pattern of small squares which could efficiently draw orienting attention in the early stages of processing (Balas and Conlin 2015). As they do not display any content and are actually meaningless, during the subsequent processing phase (P3 component) attention is withdrawn from the scrambled pictures and further processing of task-related letters is restored. On the other hand, negative pictures often relates to primary motives (such as survival and reproduction) so they are classified as important from the very beginning (Olofsson et al. 2008) and, for the same reason, their processing is maintained subsequently at the expense of the processing of letters. Thus, during the initial stage of the processing, a scrambled background may limit attention towards superimposed stimuli in the same way as a negative background. However, in the subsequent stages of processing only meaningful and important backgrounds capture attention.

These electrophysiological effects were also reflected in the one off behavioral measure. However background pictures did not change the number of errors made by participants, reaction times for letters were significantly shorter when the letters were presented against a neutral background than when they were presented against either a scrambled or negative background. Thus, the competition between the two networks visible on the electrophysiological level is confirmed by the behavioral measure. Interestingly, the results are particularly consistent with the P1 component evoked by the pictures (for which response to letters was equal for negative and scrambled pictures and lower than for neutral pictures). This suggests that the early phases of processing are the most important factor that affects reaction times. In other words, only those stimuli which attract the attention in the very early phases are capable of impairing processing of responses to concurrent stimuli.

Both types of attention constantly compete with each other even in a relatively easy task

Taken together, obtained patterns of ERPs results generally support the claim about the competition between the top-down and bottom-up attention: the more activation of the top-down attention, the less activation of the bottom-up system, and vice versa. (Berger et al. 2005, Hopfinger and West 2006, Okon-Singer et al. 2007). Moreover, our results add novel and important knowledge as to how a relatively easy task is affected by processing of task irrelevant pictures and, conversely, as to how processing of task-irrelevant pictures is affected by executing an easy task.

Firstly, the results confirm that execution of even a very simple task may deteriorate when salient pictures are presented in the background. While previous studies have only shown deterioration in relatively difficult tasks, we showed here that this deterioration is visible in a relatively easy task in which both behavioral (RT) and electrophysiological measures (ERP) indicate decreased attention towards task-relevant stimuli (Lavie 2005, MacNamara et al. 2011, Van Dillen et al. 2009). On the other hand, the study adds to the debate over automatic processing of emotionally salient yet task-irrelevant stimuli, thus confirming the recent view that processing of biologically important stimuli is automatic but still requires that minimal attentional is initiated (Pessoa et al. 2005). When one has to perform an engaging task, there are not sufficient attentional resources left for processing task-irrelevant stimuli and, as a result, differences in processing between unpleasant and neutral irrelevant pictures are no longer detectable. However, we argue that even simple tasks, such as letter recognition, may reduce automatic processing of irrelevant but emotionally-significant stimuli marked in ERP responses. Such conclusions are further justified, as processing of all pictures in our procedure (regardless of their valence) decreased when additional top-down attention influences were imposed.

Limitations

Several limitations and future directions of the study should be mentioned. Firstly, the achieved results are not entirely consistent with the hypothesis, especially those related to the N1 component. The N1 component is probably not relevant for studying competition between two attentional networks. Moreover, we hardly found any differences between the DT and UN conditions, which suggests that these tasks could similarly demand top-down attention as this is equally easy for participants. Future studies should vary the difficulty of the tasks to answer the question of whether varying difficulties are related to varying intensification of competition between networks. It would be especially valuable to check if there is a linear dependency between the competition and the difficulty of the task. Different ERP components were analyzed as markers of processing of letters (P1, P3b) and pictures (P3a, LPP). The rationale of this is mentioned in the method section and is related to the fact that different ERPs apply for processing of full screen pictures and superimposed letters and that pictures were processed more automatically than letters. Image scrambling was used to remove any meaningful content of the pictures while preserving their physical parameters like brightness or color. Since image scrambling increases high portion of the spatial spectrum, one might argue that the difference in ERPs in response to scrambled and non-scrambled pictures might be attributed to different spatial spectra of the pictures. However, the available data regarding the influence of spatial frequencies on ERPs are not conclusive (for a recent review see: Kauffmann et al. 2014). Some data suggest that only low spatial filtering might have an impact on early ERPs as those frequencies are used by magnocellular pathways (Holmes et al. 2005). As such, scrambling should not have affected ERPs in our procedure (scrambling increases high portion of the spatial spectrum). On the other hand, other data suggest that there is a linear relationship between the spectral power and the amplitude of early ERPs (P1, N1). Importantly, this relationship is observed in ERPs in response to pictures when both low and high portions of the frequency spectrum increases (De Cesarei et al. 2013).

In the light of these ambiguous data, we cannot rule out that differences in early ERPs (P1, N1) in response to scrambled and non-scrambled pictures resulted not only from the presence or absence of the meaningful content of pictures but also, at least partly, from altered spatial frequency of the pictures. As such, future studies should consider using other methods of removing meaningful content of pictures like diffeomorphic transformation (for details see: Stojanoski and Cusack 2014).

CONCLUSION

In a more ecologically valid design our data confirms the inverted relationship between two attentional networks: the more attention is volitionally directed towards task-relevant stimuli, the less is available for automatic, bottom-up processing. Conversely, the more biologically important the stimuli, the more attention they capture. Importantly, such competition might be recognizable at the electrophysiological level in a very simple task such as letter recognition.

ACKNOWLEDGMENTS

This study was supported by the Polish National Science Center (DEC-2013/09/B/HS6/02662).

The authors declare that there is no financial interest or benefit in relation to this article.

REFERENCES

Balas B, Conlin C (2015) The Visual N1 Is Sensitive to Deviations from Natural Texture Appearance. PLoS One 10: e0136471.

Berger A, Henik A, Rafal R (2005) Competition Between Endogenous and Exogenous Orienting of Visual Attention. Journal of Experimental Psychology 134: 207–221.

Clementz BA, Keil A, Kissler J (2004) Aberrant brain dynamics in schizophrenia: delayed buildup and prolonged decay of the visual steady-state response. Cognitive Brain Research 18: 121–129.

Corbetta M, Shulman GL (2002) Control of goal-directed and stimulus-driven attention in the brain. Nature Reviews Neuroscience 3: 201–215.

De Cesarei A, Mastria S, Codispoti M (2013) Early Spatial Frequency Processing of Natural Images: An ERP Study. PLoS One 8: e65103.

Delorme A, Makeig S (2004) EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics. Journal of Neuroscience 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. Biological Psychology 76: 1–10.

Dolan RJ (2002) Emotion, cognition, and behavior. Science 298: 1191–1194.

Gable PA, Harmon-Jones E (2012) Reducing attentional capture of emotion by broadening attention: increased global attention reduces early electrophysiological responses to negative stimuli. Biological Psychology 90: 150–153.

Gmaj B, Januszko P, Kamiński J, Drozdowicz E, Kopera M, Wołyńczyk-Gmaj D, Szelenberger W, Wojnar M (2016) EEG source activity during processing of neutral stimuli in subjects with anxiety disorders. Acta Neurobiologiae Experimentalis 76: 75–85.

Grzybowski S, Wyczesany M, Kaiser J (2014) The influence of context on the processing of emotional and neutral adjectives - an ERP study. Biological Psychology 99: 137–149.

Hajcak G, Dunning JP, Foti D (2009) Motivated and controlled attention to emotion: Time-course of the late positive potential. Clinical Neurophysiology 120: 505–510.

Hart SJ, Lucena N, Cleary KM, Belger A, Donkers FCL (2012) Modulation of early and late event-related potentials by emotion. Frontiers in Intergrative Neuroscience 6: 102.

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.

Hopfinger JB, West VM (2006) Interactions between endogenous and exogenous attention on cortical visual processing. NeuroImage 31: 774–789.

Jaśkiewicz M, Francuz P, Zabielska-Mendyk E, Zapała D, Augustynowicz P (2016) Effects of harmonics on aesthetic judgments of music: an ERP study involving laypersons and experts. Acta Neurobiologiae Experimentalis 76: 142–151.

Johnson R (1993) On the neural generators of the P300 component of the event-related potential. Psychophysiology, 30: 90–97.

Kauffmann L, Ramanoël S, Peyrin, C (2014) The neural bases of spatial frequency processing during scene perception. Frontiers in Intergrative Neuroscience 8.

Kossowska M, Czarnek G, Wyczesany M, Wronka E, Szwed P, Bukowski M (2015) Electrocortical indices of attention correlate with the need for closure. NeuroReport 26: 285–290.

Lavie N (2005) Distracted and confused?: selective attention under load. Trends in Cognitive Science 9: 75–82.

Luck SJ, Heinze HJ, Mangun GR, Hillyard SA (1990) Visual event-related potentials index focused attention within bilateral stimulus arrays. II. Functional dissociation of P1 and N1 components. Electroencephalography and Clinical Neurophysiology 75: 528–542.

Luck SJ, Hillyard SA, Mouloua M, Woldorff MG, Clark VP, Hawkins HL (1994) Effects of spatial cuing on luminance detectability: psychophysical and electrophysiological evidence for early selection. Journal of Experimental Psychology: Human Perception and Performance 20: 887–904.

MacNamara A, Ferri J, Hajcak G (2011) Working memory load reduces the late positive potential and this effect is attenuated with increasing anxiety. Cognitive, Affective and Behavioral Neuroscience 11: 321–331.

Marchewka A, Żurawski Ł, Jednoróg K, Grabowska A (2014) The Nencki Affective Picture System (NAPS): introduction to a novel, standardized, wide-range, high-quality, realistic picture database. Behavior Research Methods 46: 596–610.

Mayer AR, Dorflinger JM, Rao SM, Seidenberg M (2004) Neural networks underlying endogenous and exogenous visual-spatial orienting. Neuroimage 23: 534–541.

Okon-Singer H, Tzelgov J, Henik A (2007) Distinguishing between automaticity and attention in the processing of emotionally significant stimuli. Emotion 7: 147–157.

Olofsson JK, Nordin S, Sequeira H, Polich J (2008) Affective picture processing: an integrative review of ERP findings. Biological Psychology 77: 247–265.

Pessoa L, Japee S, Ungerleider LG (2005) Visual awareness and the detection of fearful faces. Emotion 5: 243–247.

Rugg MD, Milner AD, Lines CR, Phalp R (1987) Modulation of visual event-related potentials by spatial and non-spatial visual selective attention. Neuropsychologia 25: 85–96.

Stojanoski B, Cusack R (2014) Time to wave good-bye to phase scrambling: Creating controlled scrambled images using diffeomorphic transformations. Journal of Vision 14: 6–6.

Theeuwes J (2010) Top–down and bottom–up control of visual selection. Acta Psychologica 135: 77–99.

Van Dillen LF, Heslenfeld DJ, Koole SL (2009) Tuning down the emotional brain: An fMRI study of the effects of cognitive load on the processing of affective images. NeuroImage 45: 1212–1219.

Verleger R, Jaśkowski P, Wascher E (2005) Evidence for an Integrative Role of P3b in Linking Reaction to Perception. Journal of Psychophysiology 19: 165–181.

Vogel EK, Luck SJ (2000) The visual N1 component as an index of a discrimination process. Psychophysiology 37: 190–203.

Whalen PJ, Rauch SL, Etcoff NL, McInerney SC, Lee MB, Jenike MA (1998) Masked presentations of emotional facial expressions modulate amygdala activity without explicit knowledge. Journal of Neuroscience 18: 411–418.

Wronka E, Kaiser J, Coenen AML (2012) Neural generators of the auditory evoked potential components P3a and P3b. Acta Neurobiologiae Experimentalis 72: 51–64.

Wyczesany M, Ligeza TS (2017) Is elaboration of affective content sufficient to attenuate emotional responses during reappraisal procedures? Biological Psychology 127: 99–107.