

Perceptual processing of natural scenes at rapid rates: Effects of complexity, content, and emotional arousal

Andreas Löw · Margaret M. Bradley · Peter J. Lang

Published online: 19 June 2013
© Psychonomic Society, Inc. 2013

Abstract During rapid serial visual presentation (RSVP), the perceptual system is confronted with a rapidly changing array of sensory information demanding resolution. At rapid rates of presentation, previous studies have found an early (e.g., 150–280 ms) negativity over occipital sensors that is enhanced when emotional, as compared with neutral, pictures are viewed, suggesting facilitated perception. In the present study, we explored how picture composition and the presence of people in the image affect perceptual processing of pictures of natural scenes. Using RSVP, pictures that differed in perceptual composition (figure–ground or scenes), content (presence of people or not), and emotional content (emotionally arousing or neutral) were presented in a continuous stream for 330 ms each with no intertrial interval. In both subject and picture analyses, all three variables affected the amplitude of occipital negativity, with the greatest enhancement for figure–ground compositions (as compared with scenes), irrespective of content and emotional arousal, supporting an interpretation that ease of perceptual processing is associated with enhanced occipital negativity. Viewing emotional pictures prompted enhanced negativity only for pictures that depicted people, suggesting that specific features of emotionally arousing images are associated with facilitated perceptual processing, rather than all emotional content.

Keywords Emotion · Rapid serial picture presentation (RSVP) · Event-related potentials · Early negativity · Picture complexity

A. Löw (✉)
University of Greifswald,
Franz-Mehring –Str. 47, 17489 Greifswald, Germany
e-mail: andreas.loew@uni-greifswald.de

M. M. Bradley · P. J. Lang
University of Florida, Gainesville, FL, USA

During rapid serial visual presentation (RSVP) of natural scenes, perceptual processing is faced with a veritable onslaught of sensory information that demands rapid resolution. Nonetheless, a number of behavioral studies have determined that rapidly presented images are indeed perceived but quickly forgotten (Potter, 1975; Potter, Staub, Rado, & O'Connor, 2002). When perceptual processing is measured using electrophysiology during RSVP, previous studies have found that emotional, as compared with neutral, scenes prompt a larger occipito-temporal negativity starting around 150 ms after picture onset (e.g., Junghöfer, Bradley, Elbert, & Lang, 2001; Peyk, Schupp, Elbert, & Junghöfer, 2008; Schupp, Junghöfer, Weike, & Hamm, 2003), suggesting that this may be a neural signature of enhanced perceptual processing. Consistent with this, emotional pictures presented during RSVP are also recognized better on a later recognition test than are neutral pictures (Versace, Bradley, & Lang, 2010). Visual images of natural scenes, however, also vary widely in low-level perceptual properties that might also affect the ease of perceptual processing, especially when presented at rapid speeds. Among these are perceptual complexity and the presence of people. Therefore, in the present study, we assessed the extent to which complexity, content, and emotional arousal of natural scenes affected perceptual processing as measured using dense array EEG.

According to Gestalt psychologists, one of the major principles in perceptual organization is segmentation of an image into figure and ground (Palmer, 1999). When the high perceptual loads associated with the rapidly changing visual scenes in RSVP are processed, it is reasonable that images of lower perceptual complexity can be segregated more easily into coherent patterns, with these relatively simpler figure–ground pictures seeming to pop-out of the fleeting array. Thus, we presented simple figure–ground compositions, as well as more complex scenes, using RSVP and measured brain potentials, with the expectation that enhanced occipital negativity might be found for simple figure–ground scenes if

this neural event indexes enhanced perceptual processing. Moreover, we also presented emotional and neutral pictures that were either figure–ground compositions or more complex scenes, allowing us to determine the contribution of both factors to ease of perceptual processing, as indexed by occipital negativity.

A third variable that might influence perceptual processing at rapid rates is the extent to which a picture includes people. Due to their social significance (Adolphs, 1999), as well as human expertise in processing faces and people (Tarr & Gauthier, 2000), facilitated processing of faces has been reliably demonstrated (Carmel & Bentin, 2002). For instance, when faces are presented within an array of other objects, faces are quickly detected (Hershler & Hochstein, 2006), and when nonfacial stimuli need to be detected within a stimulus array, response times are prolonged when a face is present as a distracting item (Langton, Law, Burton, & Schweinberger, 2008). Moreover, when brain potentials have been measured during face processing, an early negative component (N170) that is maximal over posterior sensors is found that is similar in timing to the component modulated during RSVP, suggesting that differences between emotional and neutral stimuli might be due to the higher proportion of people in emotional, as compared with neutral, images (Colden, Bruder, & Manstead, 2008).

Thus, in the present study, using RSVP (330-ms picture duration, no intertrial interval), we presented pictures that varied in terms of perceptual composition (figure–ground, scene), presence of people (people, no people), and emotional arousal (emotional, neutral). We measured brain potentials with a focus on assessing the amplitude of occipital negativity previously found during RSVP. To the extent that specific perceptual features facilitate identification and recognition at rapid speeds, we expect that simple figure–ground compositions will prompt heightened occipital negativity, as compared with more complex scenes, and that pictures depicting people might also show more occipital negativity than pictures without people. To the extent that emotional pictures elicit enhanced processing that does not reflect perceptual or face-related features, we expect to replicate earlier findings of enhanced negativity when emotional, as compared with neutral, pictures were viewed.

Method

Participants

Twenty-seven females (mean age, 18.8 years; range, 18–21 years; 3 left-handed) with normal vision participated as part of a general psychology course requirement at the University of Florida. All participants provided written informed consent prior to participation for the study approved by the University of Florida Institutional Review Board.

Data from 2 participants were excluded from analysis due to excessive artifacts.

Materials and task

Stimuli were 128 color pictures¹ selected from the IAPS (Lang, Bradley, & Cuthbert, 2008) that covaried in terms of three variables: (1) *picture composition* (figure–ground, scene), (2) *content* (people, no people), and (3) *affective content* (emotionally arousing, neutral). Using these 128 images, eight sets (see Table 1) of 16 pictures were constructed such that the pictures in each set depicted one level of each of the three variables (e.g., figure–ground, emotional, people). Normative ratings of emotional arousal from the IAPS (Lang et al., 2008) and normative ratings of perceptual composition (0 = figure–ground, 9 = scene; Bradley, Houbova, Miccoli, Costa, & Lang, 2011) were matched across sets (see Table 1). For sets that contained emotionally arousing pictures, half of the pictures depicted pleasant contents (e.g., erotica, sports, adventure, etc.), and the other half were unpleasant pictures (e.g., mutilations, threat, violence, etc.).

Pictures were presented in blocks in which two of the eight possible sets were alternated, such that, in each block, the two sets of pictures differed for only one of the three variables. Table 2 lists the 12 possible block types resulting from all combinations of sets. For instance, one block presented pictures in a figure–ground/emotional/people set intermixed with scenes/emotional/people set, in which the pictures differed only in perceptual composition (i.e., figure–ground compositions vs. more complex scenes). In a second block type, the same pictures (i.e., figure–ground/emotional/people) were presented intermixed with pictures that differed only in content (i.e., figure–ground/emotional/no people). Finally, in yet another block, the same figure–ground/emotional/people were alternated with pictures that differed only in emotional arousal (i.e., figure–ground/neutral/people).

Across all of the blocks, each of the eight sets of pictures was presented 3 times. Within each block, 32 different stimuli (two sets of 16 pictures) were presented, with each of the 32 pictures presented 4 times with the restriction that a picture

¹ IAPS pictures used in this study were the following: figure–ground: people, emotional, 2071, 2120, 2800, 3030, 3100, 3168, 3170, 3180, 3266, 4609, 4651, 4652, 4653, 8120, 8200, 8300; neutral, 2190, 2200, 2210, 2214, 2215, 2220, 2221, 2230, 2271, 2280, 2440, 2495, 2500, 2516, 2570, 2830; no people, emotional, 1270, 5480, 5890, 6020, 6230, 6260, 6300, 7270, 7400, 7460, 8500, 8502, 8531, 9301, 9560, 9622; neutral, 6150, 7034, 7050, 7100, 7110, 7130, 7140, 7150, 7190, 7211, 7224, 7233, 7235, 7490, 7705, 7950. Scene: people, emotional, 2208, 2691, 3010, 3350, 3500, 3530, 4656, 4666, 4687, 4810, 6212, 6313, 8380, 8461, 8490, 9250; neutral, 2381, 2383, 2393, 2394, 2435, 2485, 2487, 2579, 2580, 2595, 2749, 2840, 2850, 2870, 7550, 9700; no people, emotional, 1111, 5600, 5971, 7282, 7289, 7350, 7481, 7501, 8162, 8510, 9340, 9470, 9495, 9611, 9620, 9911; neutral, 2206, 5120, 5395, 5731, 7036, 7037, 7160, 7205, 7207, 7234, 7495, 7500, 7590, 7595, 7700, 7710.

Table 1 Ratings of emotional arousal and perceptual composition for each of the eight sets of pictures in which perceptual composition, emotion, and content were covaried

Composition	Emotion	Content	Mean Arousal Rating ^a	Mean Composition Rating ^b
Figure/ground	Emotional	People	6.02	2.53
	Neutral		3.40	1.23
	Emotional	No People	5.69	1.72
	Neutral		2.92	1.48
Scene	Emotional	People	6.29	4.90
	Neutral		3.42	5.29
	Emotional	No People	5.36	5.45
	Neutral		3.40	5.42

^a 1 = low arousal; 9 = high arousal. Ratings from Lang, Bradley, & Cuthbert, 2008

^b 0 = figure-ground; 9 = scene. Ratings from Bradley et al., 2011

was repeated only once all other pictures from the set had been presented. Each picture was presented for 330 ms, immediately followed by the next picture; each block (128 trials) lasted for about 42 s, and participants passively viewed the pictures. Each of the 12 blocks (1,536 trials) was separated by 8-s interblock intervals (blank screen), and the order of presentation of the blocks was counterbalanced such that each of the 12 types (see Table 2) was presented in each serial position equally often across participants.

Data collection and analysis

The EEG was recorded continuously using a 128-channel geodesic sensor net (EGI, Eugene, OR) referenced against the vertex sensor (sampling rate, 250 Hz; bandpass, 0.1- to 100-Hz half-power cutoff frequencies; 6 dB/octave rolloff).

Impedances were kept below 30 k Ω , as recommended by the manufacturer. Data were digitally filtered offline with a low-pass of 30 Hz (zero-phase shift Butterworth filter, 12 dB/octave). Bad channels (with poor signals throughout the whole recording—e.g., due to broken sensors) were interpolated using spherical splines. On average, 4.48 channels (range: 0–8) were interpolated per recording. Ocular and cardiac artifacts were removed using BESA software (MEGIS software, Munich, Germany) as described by Ille, Berg, and Scherg (2002), and the data were then converted to the average reference. For each subject and picture set, separate averages were created. As each picture set appeared in three different blocks, data were averaged across the blocks for a specific picture set such that each of the eight averages contained the ERPs of a specific stimulus set (e.g., figure-ground/emotional/people). On average, 82.6 % of epochs were retained for analysis after exclusion of bad trials, defined as when the difference between the maximum and the minimum amplitudes on a channel exceeded a threshold of 100 μ V.

For statistical purpose, ERP data were averaged over a cluster of 19 posterior sensors (see Fig. 1). Since there is no baseline period in RSVP, occipital negativity was quantified as the difference in the mean amplitude from 116 to 140 ms post-stimulus-onset (P1) and the mean amplitude from 180 to 280 ms post-stimulus-onset. Repeated measures ANOVA of the amplitude of occipital negativity included the variables perceptual composition (figure-ground, scene), content (people, no people), and emotion (emotionally arousing, neutral). For illustration purposes, waveform plots of the data use a 0- to 48-ms post-stimulus-onset baseline.

Onset latency and topographical distributions were also compared by computing difference waves between the two levels of each factor (e.g., figure-ground minus scene). Onset latency was analyzed using a jackknife procedure in

Table 2 Description of the 12 different types of blocks that result when two sets of pictures are presented in each block that differ on only one of the varied factors (i.e., composition, content, emotion)

Block	Set 1	Set 2
1	Figure/Ground - Emotional - People	Scene - Emotional - People
2	Figure/Ground - Neutral - People	Scene - Neutral - People
3	Figure/Ground - Emotional - No People	Scene - Emotional - No People
4	Figure/Ground - Neutral - No People	Scene - Neutral - No People
5	Figure/Ground - Emotional - People	Figure/Ground - Emotional - No People
6	Figure/Ground - Neutral - People	Figure/Ground - Neutral - No People
7	Scene - Emotional - People	Scene - Emotional - No People
8	Scene - Neutral - People	Scene - Neutral - No People
9	Figure/Ground - Emotional - People	Figure/Ground - Neutral - People
10	Figure/Ground - Emotional - No People	Figure/Ground - Neutral - No People
11	Scene - Emotional - People	Scene - Neutral - People
12	Scene - Emotional - No People	Scene - Neutral - No People

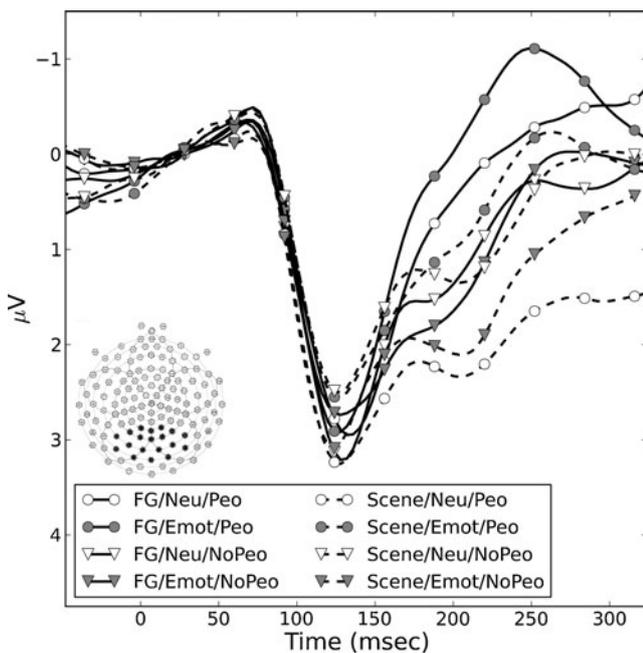


Fig. 1 ERP waveforms (averaged across a cluster of 19 sensors highlighted in the sensor layout) during RSVP for each of the eight different picture sets formed by covarying composition (figure–ground [FG], solid lines), scene, dashed lines), emotion (emotional [Emot], filled; neutral [Neu], open), and content (people [Peo], circles; no people [NoPeo], triangle). The most pronounced negativity occurs for figure–ground compositions that depict people in emotional contexts

which 25 grand averages were computed, each time omitting the data from 1 participant. Onset latency was determined for each difference waveform as the point at which 50 % of the peak amplitude was reached (for details, see Kiesel, Miller, Jolicoeur, & Brisson, 2008; Luck et al., 2009), and an ANOVA was used to assess effects of each of the three factors, with adjustment of the degrees of freedom (Kiesel et al., 2008).

Topographical maps of the difference potentials were compared pairwise by computing global map dissimilarity (Lehmann & Skrandies, 1980), defined as the root mean square of the difference between two normalized (divided by global field power) maps. This index varies between 0 (maps have identical voltage distribution) and 2 (maps have inverse voltage distributions). Statistical significance of the global map dissimilarity values was determined by nonparametric randomization tests, in which an empirical distribution of global map dissimilarity values was created by randomly assigning the single-subject maps to experimental conditions and calculating the global map dissimilarity value for the resulting group-average ERPs (10,000 permutations). The value from the actual group-average ERPs was then compared with this distribution (see Murray, Brunet, & Michel, 2008, for details). Topographies were tested at each of the time points in which the global field power of the difference wave was maximal.

Pictures analysis

An analysis focusing on ERPs as a function of picture, rather than participant, was undertaken by creating a mean waveform across the 12 presentations of each of the 128 pictures for each participant and then averaging across participants, resulting in an evoked potential for each picture. This picture analysis assessed the influence of the manipulated dimensions on occipital negativity using regression analysis. In these analyses, ratings of emotional arousal (from the IAPS; Lang et al., 2008) were used to index the emotionality of each picture; previous studies have determined that a number of measures of emotional engagement, including skin conductance, pupil dilation, the amplitude of a centro-parietal late positive potential, and BOLD changes in visual sensory and limbic cortex covary with ratings of emotional arousal (see Lang & Bradley, 2010, for an overview), regardless of hedonic valence (e.g., pleasantness). Moreover, pictures rated lowest in emotional arousal are those that are rated as neutral in hedonic valence, allowing these ratings to continuously code the emotional arousal of a specific picture (Lang, 2010). Relatedly, normative ratings of perceptual composition (Bradley et al., 2011) were used to continuously index the perceptual composition of each picture; content (people, no people) was coded as a nominal variable in these item analyses.

Results

Figures 1 and 2 illustrate the RSVP waveforms averaged across occipital sensors. Occipital negativity was significantly modulated by each of the three manipulated variables: Pictures with figure–ground composition evoked more negative-going potentials than did complex scenes [composition: $F(1, 24) = 85.91, p < .001, \eta_p^2 = .78$; see Fig. 2, left], and pictures depicting people elicited enhanced negativity, as compared with pictures that did not include people [content: $F(1, 24) = 46.78, p < .001, \eta_p^2 = .66$; see Fig. 2, middle]. Replicating previous research, emotional, as compared with neutral, pictures also elicited enhanced negativity [emotion: $F(1, 24) = 8.38, p < .01, \eta_p^2 = .26$; see Fig. 2, right].

A number of significant two-way interactions qualified these main effects. A significant interaction of composition and content, $F(1, 24) = 21.38, p < .001, \eta_p^2 = .47$, was followed up by simple main effects tests that indicated that occipital negativity was larger for figure–ground pictures, as compared with scenes, regardless of whether pictures depicted people [composition: $F(1, 24) = 106, p < .001$] or not [composition: $F(1, 24) = 14.4, p < .005$]. On the other hand, the presence of people in a picture evoked more negativity than did those without people only for simple figure–ground compositions [content, $F(1, 24) = 42.20, p < .001$]. Relatedly, although the interaction of composition and emotion was only

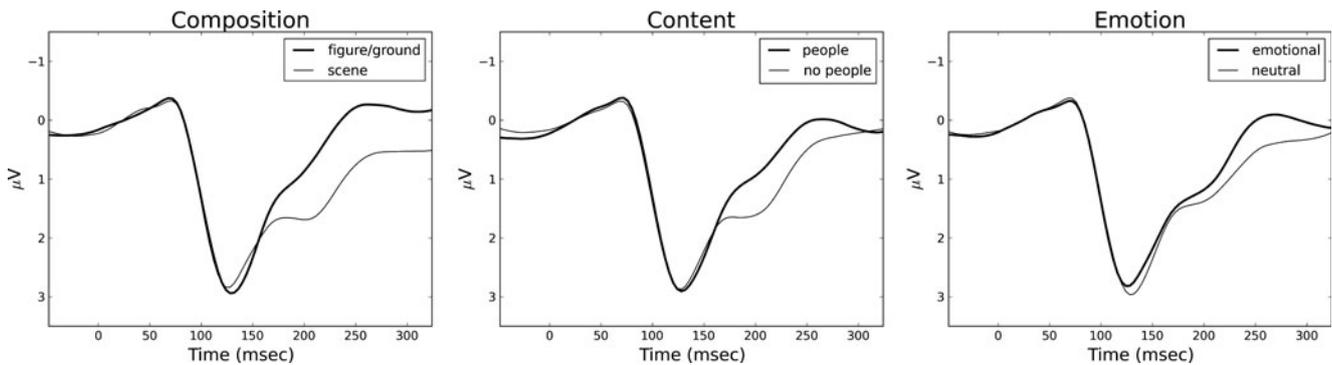


Fig. 2 Averaged ERP waveforms illustrating the main effects of varying perceptual composition (figure–ground and scene), content (people and no people), and emotion (emotional and neutral pictures) during RSVP

marginal, $F(1, 24) = 3.72, p < .07$ (see Table 3), simple main effects again indicated that occipital negativity was larger for figure–ground pictures, as compared with scenes, regardless of whether pictures were emotionally arousing [composition: $F(1, 24) = 46, p < .001$] or neutral [composition: $F(1, 24) = 87, p < .001$]. On the other hand, for simple figure–ground compositions, in which occipital negativity was already quite enhanced, emotional content did not further modulate negativity, $F < 1$, while for scenes, negativity was enhanced for emotional, as compared with neutral, pictures [emotion: $F(1, 24) = 18.9, p < .001$].

An interaction of content and emotion, $F(1, 24) = 27.97, p < .001, \eta_p^2 = .54$, indicated that emotional pictures elicited enhanced negativity, as compared with neutral pictures for pictures that included people [emotion: $F(1, 24) = 25.14, p < .001$], whereas the difference was in the opposite direction for pictures that did not include people [emotion: $F(1, 24) = 9.38, p < .005$]. Moreover, whereas there was no difference in negativity whether neutral pictures depicted people or not, $F < 1$, pictures of emotional people elicited enhanced negativity, as compared with emotional pictures that did not include people [content: $F(1, 24) = 46.18, p < .001$].

Table 3 Mean amplitude (μV) of early occipital negativity for the interaction effects (with standard errors in parentheses)

	Composition \times Content	
	Figure–ground	Scene
People	–3.04 (0.40)	–1.62 (0.36)
No people	–2.04 (0.37)	–1.48 (0.35)
	Emotion \times Composition	
	Figure–ground	Scene
Emotional	–2.59 (0.39)	–1.73 (0.36)
Neutral	–2.48 (0.38)	–1.38 (0.36)
	Emotion \times Content	
	Emotional	Neutral
People	–2.71 (0.40)	–1.95 (0.37)
No people	–1.61 (0.35)	–1.91 (0.36)

Onset latency

Average onset latency varied across conditions, $F_{\text{corr}}(2, 48) = 7.23, p < .003$, with an estimated onset of 184 ms for differences due to perceptual composition, 175 ms for differences due to content (i.e., presence of people), and 212 ms for emotional differences. Pairwise comparisons indicated that onset latency as a function of perceptual composition and the presence of people did not differ from each other (composition vs. content: $p > .28$) but that both of these differences began earlier than effects due to emotional arousal (composition vs. emotion, $p < .03$; content vs. emotion, $p < .002$).

Topographies

Figure 3 illustrates the similar topography of occipital negativity in the difference scalp maps for composition, content, and emotional arousal, as well as in bilateral occipital current sources in the current source density maps. Pairwise comparisons of the topographical distributions using global map dissimilarity (DISS) indicated that there were no significant differences in spatial voltage distribution as a function of picture composition, content, or emotion (composition vs. content, DISS = .50, $p = .12$; composition vs. emotion, DISS = .33, $p = .35$; content vs. emotion, DISS = .45, $p = .19$).

Item analysis

In the picture analysis, stepwise regression (in which effects due to composition, content, or emotion were entered on the basis of F values) agreed with the subject analyses in indicating that each of the three manipulated variables accounted for significant variance in occipital negativity. Perceptual composition accounted for the most variance (21 %), $F(1, 126) = 33.57, p < .001$, followed by content (6.7 %), $F(1, 125) = 11.66, p < .001$, and then emotional arousal (4.2 %), $F(1, 124) = 7.84, p < .01$. Figure 4 illustrates these relationships: When pictures are ordered by rated composition (top panel), occipital negativity is enhanced

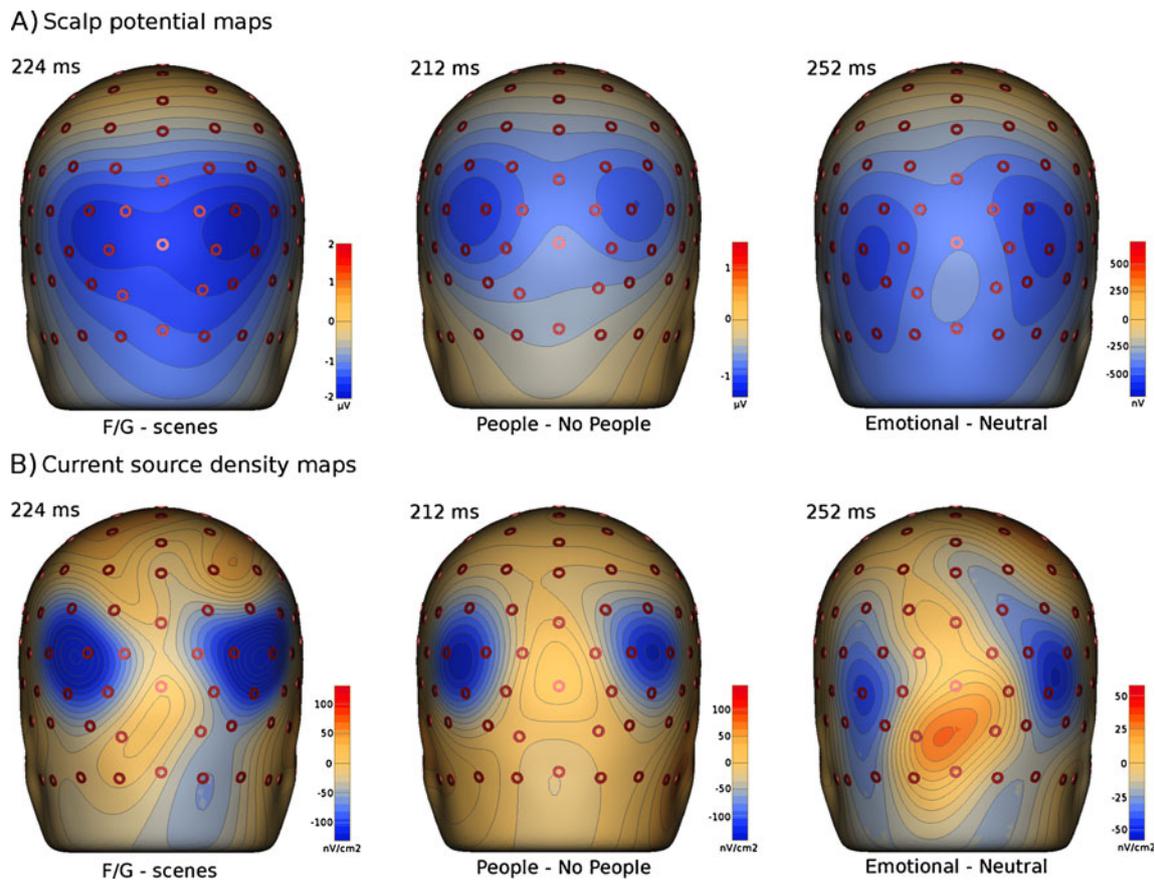


Fig. 3 **a** Scalp potentials maps (*back view*) for the difference between the two levels of each factor (composition, left; content, middle; emotion, right) at the time point where global field power was largest. **b** Current source density maps (*back view*) of the difference scalp potentials shown in Panel A

for simple figure–ground pictures, as compared with complex scenes, regardless of whether pictures are emotional (top left) or neutral (top right), with slightly stronger relationships for pictures that include people. When pictures are ordered by rated emotional arousal (bottom panel), on the other hand, the most arousing pictures are associated with the most enhanced negativity for pictures that include people (bottom left), whether these are simple figure–ground compositions or more complex scenes. For pictures that do not include people, there are no systematic differences in occipital negativity as a function of rated emotional arousal (bottom right).

Discussion

During rapid serial visual presentation, specific properties of natural scenes facilitated perceptual processing as indexed by enhanced negativity (180–280 ms) over occipital sensors. Perceptually simple figure–ground compositions were associated with enhanced occipital negativity, as compared with more complex scenes, as were pictures depicting people, as compared with those that did not, and pictures of emotionally arousing, as

compared with neutral, scenes. When occipital negativity was computed individually for each picture presented using RSVP, regression analysis provided further information regarding the relative impact of specific features in facilitating recognition at rapid rates, with perceptual composition accounting for the most variance and emotional arousal accounting for the least. Rather than suggesting that processing any picture with emotional content is facilitated, the data instead suggest that specific perceptual features in images that depict people in pleasant and aversive situations are associated with enhanced RSVP processing. Because these pictures typically include multiple exemplars of erotica and violence, it is likely that features related to naked and/or injured body parts may mediate perceptual facilitation, as indexed by occipital negativity.

Enhanced negativity was found for figure–ground compositions, as compared with more complex scenes, regardless of whether these depicted emotional or neutral content and whether they included people or not, confirming that figure–ground composition is a key variable facilitating the ease of perceptual processing during RSVP of natural scenes. And, consistent with hypotheses and data showing that humans are experts at face/person detection (Bentin, Allison, Puce, Perez, & McCarthy, 1996; Kanwisher &

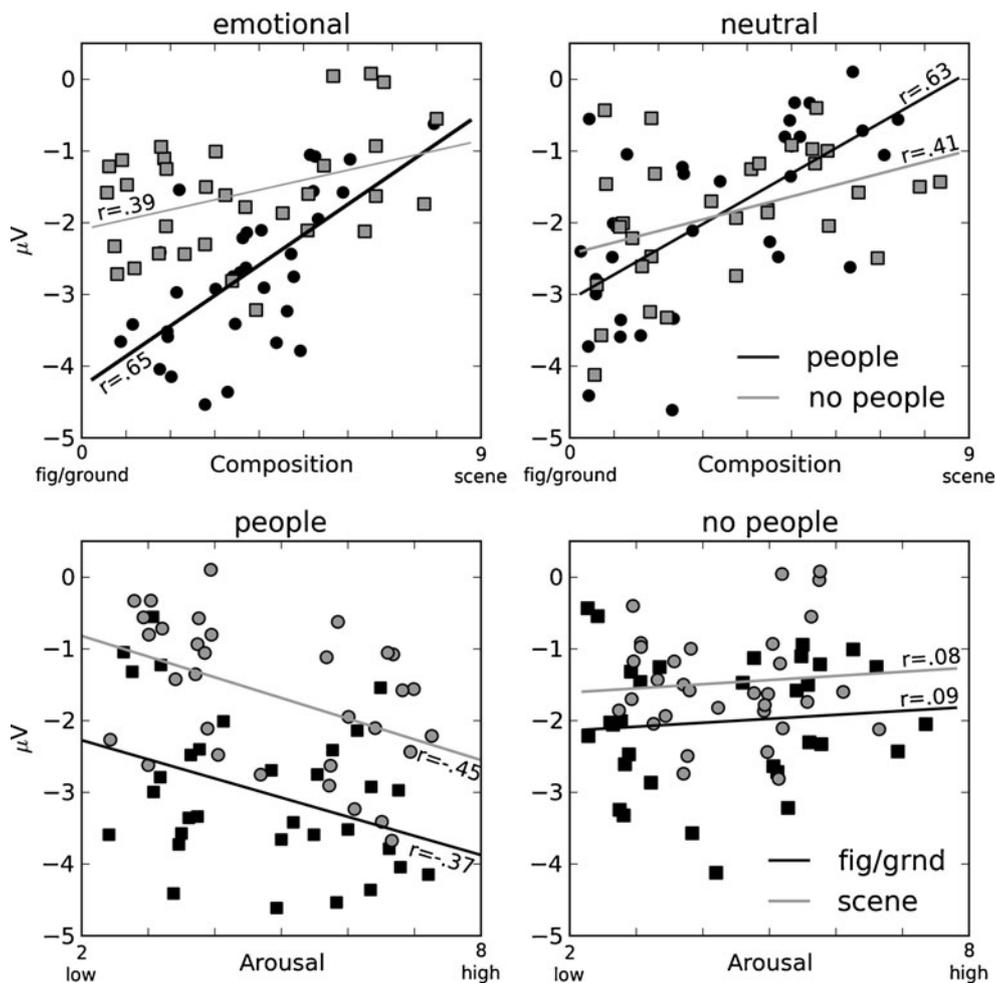


Fig. 4 Picture analysis. Top panel: Scatterplots of the relationship between amplitude of early occipital negativity and rated perceptual composition of emotional (left) and neutral (right) pictures for people (black circles) and objects (gray squares). The lines represent the linear regression for people (black lines) and objects (gray lines). Bottom

panel: Scatterplots of the relationship between amplitude of the early occipital negativity and rated arousal of figure–ground pictures (black circles) and scenes (gray squares) for people (left) and objects (right). The lines represent the linear regression for figure–ground pictures (black lines) and scenes (gray lines)

Yovel, 2006), probably mediated by both experience and social relevance, simple figure–ground pictures that depicted people were associated with enhanced perceptual processing, as compared with pictures that did not include people.

A number of studies have reported an occipital-temporal negativity that is maximal around 170 ms following the onset of simple faces, as compared with objects (e.g., Bentin et al., 1996; Itier & Taylor, 2004), raising the issue of whether perceptual facilitation during RSVP of natural scenes primarily reflects enhanced face processing. As compared with the potentials measured during RSVP, however, the typical N170 has a more temporal topography and a much clearer peak (perhaps resulting from the typically slower presentation rate; Rossion & Jacques, 2008). Moreover, in the present study, pictures without people also prompted enhanced occipital negativity if these were simple figure–ground perceptual compositions, and the presence of people did not facilitate processing of complex scenes, making it difficult to

attribute differences in perceptual processing during RSVP solely to the mere presence of people in a picture.

Whether a picture depicted emotional or neutral scenes controlled the least amount of variance in occipital negativity and was constrained to specific contexts, with the largest negativity found for simple figure–ground pictures that included people in emotional contexts. Previous studies have clearly determined that emotional engagement when pictures of natural scenes are viewed is strongest for pictures that depict people in sexual, violent, and other highly arousing emotional contexts (Bradley, Codispoti, Cuthbert, & Lang, 2001; Lang & Bradley, 2010; Schupp et al., 2004; Weinberg & Hajcak, 2010) and, furthermore, that an electrophysiological index of enhanced processing, the late positive potential, is of greatest amplitude when emotionally arousing pictures that are simple figure–ground compositions are viewed (Bradley, Hamby, Löw, & Lang, 2007; Nordström & Wiens, 2012). Perceptual enhancement for highly arousing emotional scenes during RSVP

suggests that, in addition to perceptual composition and the presence of faces, some emotional features show perceptual facilitation.

It is of course possible that this subset of pictures also differs along some unspecified physical dimension(s), which remains an avenue for future investigation. Indeed, quantifying and equating information in emotional and neutral images is challenging not only because the variables potentially affecting perceptual processing are many, but also because not all of the critical factors are yet understood. Previous studies assessing effects of specific picture content on early brain potentials, however, have reported that pictures of erotica prompt the largest (Schupp et al., 2007; Weinberg & Hajcak, 2010) early enhancement, when compared with other affective contents. Assuming that effects of emotional content on early perceptual identification are mediated by the presence of specific perceptual features that activate representations linked to the subcortical appetitive and defensive systems that direct emotional reactivity (Lang & Bradley, 2010), one task for future research is to identify these features and determine their contribution to perceptual processing.

Moreover, whereas perceptual composition reliably affects perceptual processing at both fast and slow rates of presentation (Bradley et al., 2007; Wiens, Sand, & Olofsson, 2011), effects of emotion are more variable. Thus, whereas emotional content affected occipital negativity in the present RSVP study, emotion had no effect on these brain potentials in a study using a slower (6 s) presentation rate in which perceptual composition was again experimentally controlled (Bradley et al., 2007). Finding that variables differentially modulate brain potentials as a function of presentation rate is not surprising, given the sheer amount and variety of information presented per unit of time during RSVP, which poses challenges for perceptual processing and, presumably, prompts overlapping (as well as offset) potentials. And, although specific effects of emotional content on the amplitude of occipital negativity were found during RSVP in the present study, latency analyses suggested that this facilitation may have a later onset (but same topographical distribution), as compared with effects due to perceptual composition or the presence of people. One interpretation of this finding is that the features facilitating perceptual processing for emotional images may not be as fundamental as figure–ground segmentation or associated with very high expertise and familiarity in processing human faces but, instead, require more time for resolution.

Taken together, the present study illustrates a multifactorial approach to determining the contribution of specific factors to visual perception of natural scenes. When covarying perceptual composition, content, and emotion of natural scenes presented using RSVP, we found that all three variables contribute to the amplitude of occipital negativity (180–280 ms), supporting the inference that ease of perceptual processing at rapid rates is facilitated by figure–ground composition, the

presence of people, and, to a lesser extent, emotional content. The data are consistent with a visual-processing stream for natural scenes in which figure–ground segmentation and the presence of people show the earliest facilitation in perceptual processing, with specific features in pictures of people in emotional contexts facilitating perception somewhat later. From a conceptual viewpoint, the data suggest that early perceptual processing of emotional images probably does not reflect enhanced selective attention to emotional cues in general but, rather, that some perceptual features associated with specific emotional contexts are perceived more easily, reflecting familiarity, priming, or, possibly, preparedness. Identifying these features and the mechanism of their action remains for future research.

Author note This research was supported in part by NIMH R01 MH094386.

References

- Adolphs, R. (1999). Social cognition and the human brain. *Trends in Cognitive Sciences*, 3(12), 469–479.
- Bentin, S., Allison, T., Puce, A., Perez, E., & McCarthy, G. (1996). Electrophysiological Studies of Face Perception in Humans. *Journal of Cognitive Neuroscience*, 8(6), 551–565. doi:10.1162/jocn.1996.8.6.551
- Bradley, M. M., Codispoti, M., Cuthbert, B. N., & Lang, P. J. (2001). Emotion and motivation I: Defensive and appetitive reactions in picture processing. *Emotion (Washington, D.C.)*, 1(3), 276–298. doi:10.1037//1528-3542.1.3.276
- Bradley, M. M., Hamby, S., Löw, A., & Lang, P. J. (2007). Brain potentials in perception: Picture complexity and emotional arousal. *Psychophysiology*, 44(3), 364–373. doi:10.1111/j.1469-8986.2007.00520.x
- Bradley, M. M., Houbova, P., Miccoli, L., Costa, V. D., & Lang, P. J. (2011). Scan patterns when viewing natural scenes: Emotion, complexity, and repetition. *Psychophysiology*, 48(11), 1544–1553. doi:10.1111/j.1469-8986.2011.01223.x
- Carmel, D., & Bentin, S. (2002). Domain specificity versus expertise: Factors influencing distinct processing of faces. *Cognition*, 83(1), 1–29.
- Colden, A., Bruder, M., & Manstead, A. S. R. (2008). Human content in affect-inducing stimuli: A secondary analysis of the international affective picture system. *Motivation and Emotion*, 32(4), 260–269. doi:10.1007/s11031-008-9107-z
- Hershler, O., & Hochstein, S. (2006). With a careful look: Still no low-level confound to face pop-out. *Vision Research*, 46(18), 3028–3035. doi:10.1016/j.visres.2006.03.023
- Ille, N., Berg, P., & Scherg, M. (2002). Artifact correction of the ongoing EEG using spatial filters based on artifact and brain signal topographies. *Journal of Clinical Neurophysiology*, 19(2), 113–124.
- Itier, R. J., & Taylor, M. J. (2004). Source analysis of the N170 to faces and objects. *Neuroreport*, 15(8), 1261–1265. doi:10.1097/01.wnr.0000127827
- Junghöfer, M., Bradley, M. M., Elbert, T. R., & Lang, P. J. (2001). Fleeting images: A new look at early emotion discrimination. *Psychophysiology*, 38(2), 175–178.
- Kanwisher, N., & Yovel, G. (2006). The fusiform face area: A cortical region specialized for the perception of faces. *Philosophical*

- Transactions of the Royal Society of London. Series B, Biological Sciences*, 361(1476), 2109–2128. doi:10.1098/rstb.2006.1934
- Kiesel, A., Miller, J., Jolicoeur, P., & Brisson, B. (2008). Measurement of ERP latency differences: A comparison of single-participant and jackknife-based scoring methods. *Psychophysiology*, 45(2), 250–274. doi:10.1111/j.1469-8986.2007.00618.x
- Lang, P. J. (2010). Emotion and motivation: Toward consensus definitions and a common research purpose. *Emotion Review*, 2(3), 229–233. doi:10.1177/1754073910361984
- Lang, P. J., & Bradley, M. M. (2010). Emotion and the motivational brain. *Biological Psychology*, 84(3), 437–450. doi:10.1016/j.biopsycho.2009.10.007
- Lang, Peter J., Bradley, M. M., & Cuthbert, B. N. (2008). *International affective picture system (IAPS): Affective ratings of pictures and instruction manual. Technical Report A-7*. Gainesville, FL.
- Langton, S. R. H., Law, A. S., Burton, A. M., & Schweinberger, S. R. (2008). Attention capture by faces. *Cognition*, 107(1), 330–342. doi:10.1016/j.cognition.2007.07.012
- Lehmann, D., & Skrandies, W. (1980). Reference-free identification of components of checkerboard-evoked multichannel potential fields. *Electroencephalography and Clinical Neurophysiology*, 48(6), 609–621.
- Luck, S. J., Kappenman, E. S., Fuller, R. L., Robinson, B., Summerfelt, A., & Gold, J. M. (2009). Impaired response selection in schizophrenia: Evidence from the P3 wave and the lateralized readiness potential. *Psychophysiology*, 46(4), 776–786. doi:10.1111/j.1469-8986.2009.00817.x
- Murray, M. M., Brunet, D., & Michel, C. M. (2008). Topographic ERP analyses: A step-by-step tutorial review. *Brain Topography*, 20(4), 249–264. doi:10.1007/s10548-008-0054-5
- Nordström, H., & Wiens, S. (2012). Emotional event-related potentials are larger to figures than scenes but are similarly reduced by inattention. *BMC Neuroscience*, 13, 49. doi:10.1186/1471-2202-13-49
- Palmer, S. E. (1999). *Vision Science: Photons to Phenomenology*. The MIT Press.
- Peyk, P., Schupp, H. T., Elbert, T., & Junghöfer, M. (2008). Emotion processing in the visual brain: A MEG analysis. *Brain Topography*, 20(4), 205–215. doi:10.1007/s10548-008-0052-7
- Potter, M. C. (1975). Meaning in visual search. *Science (New York, N.Y.)*, 187(4180), 965–966.
- Potter, M. C., Staub, A., Rado, J., & O'Connor, D. H. (2002). Recognition memory for briefly presented pictures: The time course of rapid forgetting. *Journal of Experimental Psychology: Human Perception and Performance*, 28(5), 1163–1175.
- Rossion, B., & Jacques, C. (2008). Does physical interstimulus variance account for early electrophysiological face sensitive responses in the human brain? Ten lessons on the N170. *NeuroImage*, 39(4), 1959–1979. doi:10.1016/j.neuroimage.2007.10.011
- Schupp, H., Cuthbert, B., Bradley, M., Hillman, C., Hamm, A., & Lang, P. (2004). Brain processes in emotional perception: Motivated attention. *Cognition & Emotion*, 18(5), 593–611. doi:10.1080/02699930341000239
- Schupp, H. T., Junghöfer, M., Weike, A. I., & Hamm, A. O. (2003). Attention and emotion: An ERP analysis of facilitated emotional stimulus processing. *Neuroreport*, 14(8), 1107–1110. doi:10.1097/01.wnr.0000075416.59944.49
- Schupp, H. T., Stockburger, J., Codispoti, M., Junghöfer, M., Weike, A. I., & Hamm, A. O. (2007). Selective visual attention to emotion. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 27(5), 1082–1089. doi:10.1523/JNEUROSCI.3223-06.2007
- Tarr, M. J., & Gauthier, I. (2000). FFA: A flexible fusiform area for subordinate-level visual processing automatized by expertise. *Nature Neuroscience*, 3(8), 764–769. doi:10.1038/77666
- Versace, F., Bradley, M. M., & Lang, P. J. (2010). Memory and event-related potentials for rapidly presented emotional pictures. *Experimental Brain Research*, 205(2), 223–233. doi:10.1007/s00221-010-2356-6
- Weinberg, A., & Hajcak, G. (2010). Beyond good and evil: The time-course of neural activity elicited by specific picture content. *Emotion (Washington, D.C.)*, 10(6), 767–782. doi:10.1037/a0020242
- Wiens, S., Sand, A., & Olofsson, J. K. (2011). Nonemotional features suppress early and enhance late emotional electrocortical responses to negative pictures. *Biological Psychology*, 86(1), 83–89. doi:10.1016/j.biopsycho.2010.11.001