Rapid extraction of emotion regularities from complex scenes in the human brain

Supplementary materials

Antonio Schettino[†a,b], Christopher Gundlach[†c], & Matthias M. Müller[c*]

[a] Department of Experimental-Clinical and Health Psychology, Ghent University, 9000 Ghent, Belgium

[b] Institute for Globally Distributed Open Research and Education (IGDORE)

[c] Institute of Psychology, University of Leipzig, 04109 Leipzig, Germany

[†] Co-first authors

* Corresponding author:

Matthias M. Müller

Institute of Psychology

University of Leipzig

Neumarkt 9-19 - 04109 Leipzig (Germany)

Phone: +49 (0) 341 97 39 543

Email: m.mueller@rz.uni-leipzig.de

Stimuli

The pictures selected from the *IAPS* (Lang, Bradley, & Cuthbert, 2008) and *EmoPics* (Wessa et al., 2010) databases (see *Appendix*) depict humans engaged in various activities (e.g., sport, sexual intercourse, surgical operations), threatening and innocuous animals (e.g., dogs, snakes, turtles), and inanimate objects (e.g., ships, money, rockets). These visual stimuli were divided into three emotion categories – neutral, unpleasant, and pleasant – according to their normative valence and arousal ratings (see *Supplementary Table S1*). Below we statistically verify the efficacy of our preselection.

Table S1. Descriptive statistics of the pictures selected for the experiment: number of stimuli displaying living/nonliving objects; normative valence, arousal, and complexity ratings; apparent contrast and jpeg size; valence and arousal ratings obtained after the main SSVEP experiment.

emotion class	living/ nonliving	normative valence	normative arousal	normative complexity	apparent contrast	jpeg size	ratings valence	ratings arousal
neutral	37/13	5.01 *(0.13)*	4.54 *(0.08)*	4.38 *(0.14)*	0.68 *(0.04)*	105 *(2.70)*	5.04 *(0.05)*	3.44 *(0.32)*
pleasant	43/7	6.94 *(0.08)*	6.44 *(0.06)*	4.48 *(0.18)*	0.60 *(0.03)*	100 *(2.78)*	6.23 *(0.13)*	5.53 *(0.15)*
unpleasant	41/9	2.50 *(0.07)*	6.41 *(0.06)*	4.55 *(0.19)*	0.69 *(0.04)*	103 *(2.51)*	2.46 *(0.14)*	6.04 *(0.27)*

Note. 20% trimmed means and standard errors (in parentheses). Arousal, valence, and complexity ratings range from 1 to 9, with 9 being the highest arousal, the most positive valence, and the highest subjective complexity. Apparent contrast ranges from 0 to 1, and jpeg size is expressed in kilobytes.

Valence and arousal

A one-way bootstrap rbANOVA (5,000 samples) on **valence** ratings showed a significant effect of emotion (F_t = 912, *p* < .001, ξ = .99, $CI_{95\%}$ [.90, 1.00]). Post-hoc tests (Field & Wilcox, 2017) showed higher ratings for pleasant relative to neutral (ψ = -1.93, *p* < .001, *g* = 3.41, $CI_{95\%}$ [2.91, 3.91]) and unpleasant scenes (ψ = 4.44, *p* < .001, *g* = -8.72, $CI_{95\%}$ [-10.20, -7.32]), as well as for neutral compared to unpleasant scenes (ψ = 2.52, *p* < .001, *g* = -4.16, $CI_{95\%}$ [-4.85, -3.45]). A Bayesian ANOVA confirmed that the data favored the model with the effect of emotion (H_1)

5.97 x 10^{75} times more than the null model (H_0) (see *Supplementary Table S2*). Bayesian *t*-tests confirmed higher ratings for pleasant relative to neutral (BF_{10} = 6.21 x 10^{27} ±0.00%) and unpleasant scenes (BF_{10} = 2.01 x 10^{62} ±0.00%), as well as for neutral vs. unpleasant scenes (BF_{10} = 2.21 x 10^{34} ±0.00%).

A one-way bootstrap rbANOVA on **arousal** ratings showed a significant effect of emotion (F_t = 194, p < .001, ξ = .95, $CI_{95\%}$ [.93, .98]). Post-hoc tests showed lower arousal for neutral compared to pleasant (ψ = -1.90, p < .001, g = 4.31, $CI_{95\%}$ [3.55, 5.03]) and unpleasant scenes (ψ = -1.87, p < .001, g = 4.50, $CI_{95\%}$ [3.65, 5.29]), whereas pleasant and unpleasant scenes did not statistically differ (ψ = 0.03, p = .695, g = -0.16, $CI_{95\%}$ [-0.54, 0.23]). A Bayesian ANOVA confirmed that the data favored H_1 2.43 x 10^{54} times more than H_0. Bayesian *t*-tests confirmed lower arousal ratings for neutral relative to pleasant (BF_{10} = 4.35 x 10^{35} ±0.00%) and unpleasant scenes (BF_{10} = 1.37 x 10^{37} ±0.00%), whereas H_0 ought to be preferred when comparing ratings of pleasant and unpleasant stimuli (BF_{10} = 0.29 ±0.00%).

Other (visual) properties

Given the rapid stimulus presentation employed in our study, it was important to verify that our stimulus classes did not differ with respect to other visual properties, which could be prioritized over emotional content because they may be more easily processed. We controlled for apparent contrast (i.e., standard deviation divided by mean pixel luminance) and two measures of picture complexity: ratings collected in a sample of undergraduates in the context of earlier studies (Bradley, Hamby, Löw, & Lang, 2007; Schettino, Keil, Porcu, & Müller, 2016) and jpeg size (Bates et al., 2003; Junghöfer, Bradley, Elbert, & Lang, 2001; Marin & Leder, 2013; Schettino et al., 2016).

A one-way bootstrap rbANOVA on **apparent contrast** showed no significant effect of emotion (F_t = 1.72, p = .197, ξ = .24, $CI_{95\%}$ [.00, .51]). Post-hoc tests confirmed no significant differences between pleasant and neutral (ψ = 0.08, p = .129, g = -0.25, $CI_{95\%}$ [-0.64, 0.16]), pleasant and unpleasant (ψ = -0.09, p = .089, g = 0.27, $CI_{95\%}$ [-0.14, 0.66]), and neutral and unpleasant scenes (ψ = -0.01, p = .831, g = 0.02, $CI_{95\%}$ [-0.38, 0.41]). A Bayesian ANOVA showed that the data favored H_0 1 / 0.10 ≈ 10 times more than H_1. Bayesian *t*-tests confirmed that all differences were more likely under H_0: pleasant vs. neutral (BF_{10} = 0.42 ±0.00%), pleasant vs. unpleasant (BF_{10} = 0.48 ±0.00%), and neutral vs. unpleasant scenes (BF_{10} = 0.21 ±0.03%).

A one-way bootstrap rbANOVA on **complexity** ratings showed no significant effect of emotion (F_t = 0.28, p = .751, ξ = .07, $CI_{95\%}$ [.00, .39]). Post-hoc tests confirmed no significant differences between pleasant and neutral (ψ = 0.08, p = .129, g = -0.25, $CI_{95\%}$ [-0.64, 0.16]), pleasant and unpleasant (ψ = -0.09, p = .09, g = 0.27, $CI_{95\%}$ [-0.14, 0.66]), and neutral and unpleasant scenes (ψ = -0.01, p = .831, g = 0.02, $CI_{95\%}$ [-0.38, 0.41]). A Bayesian ANOVA showed that the data favored H_0 1 / 0.04 ≈ 25 times more than H_1. Bayesian *t*-tests confirmed that all differences were more likely under H_0: pleasant and neutral (BF_{10} = 0.21 ±0.03%), pleasant vs. unpleasant (BF_{10} = 0.22 ±0.03%), and neutral vs. unpleasant scenes (BF_{10} = 0.21 ±0.03%).

Finally, a one-way bootstrap rbANOVA on **jpeg size** showed no significant effect of emotion (F_t = 0.66, p = .520, ξ = .16, $CI_{95\%}$ [.00, .44]). Post-hoc tests confirmed no significant differences between pleasant and neutral (ψ = 4.46, p = .320, g = -0.14, $CI_{95\%}$ [-0.55, 0.25]), pleasant and unpleasant (ψ = -2.69, p = .503, g = 0.13, $CI_{95\%}$ [-0.27, 0.53]), and neutral and unpleasant scenes (ψ = 1.77, p = .696, g = -0.02, $CI_{95\%}$ [-0.40, 0.38]). A Bayesian ANOVA showed that the data favored H_0 1 / 0.05 ≈ 20 times more than H_1. Bayesian *t*-tests confirmed that all differences were more likely under H_0: pleasant vs. neutral (BF_{10} = 0.27 ±0.03%), pleasant vs. unpleasant (BF_{10} = 0.25 ±0.03%), and neutral vs. unpleasant scenes (BF_{10} = 0.21 ±0.03%).

Animacy

Previous studies have shown that humans' visual attention tends to be more rapidly oriented towards living creatures compared to inanimate objects (New, Cosmides, & Tooby, 2007; Proverbio, Del Zotto, & Zani, 2007; Thorpe, Fize, & Marlot, 1996). To avoid that such animacy effects would confound any hypothesized effects of emotional attention, we attempted to balance the proportion of pictures displaying animate and inanimate objects across the three emotion classes (see *Supplementary Table S1*). A chi-square test of independence showed no significant differences ($\chi^2_{2,\,N\,=\,150}$ = 2.39, p = .302, Cramer's V = .13, $CI_{95\%}$ [0.00, 0.28]), further confirmed by the independent multinomial Gûnel and Dickey BF^{I}_{10} = 0.10 ±0.00% (Gûnel & Dickey, 1974; Jamil et al., 2016), providing support for the hypothesis that there are no systematic differences in the number of pictures displaying animate and inanimate objects across emotion categories.

Table S2. Bayes factors (BF_{10}) and percentage of proportional errors (% pe) for the emotion compared to the null model as well as each paired or independent comparison, obtained by using JZS priors with different scaling factors.

dependent variable	test	*r* = 1		*r* = .707		*r* = .5	
		BF_{10}	% pe	BF_{10}	% pe	BF_{10}	% pe
normative valence	ANOVA	8.25×10^{75}	±0.00	5.97×10^{75}	±0.00	4.27×10^{75}	±0.00
	N vs. P	8.43×10^{27}	±0.00	6.21×10^{27}	±0.00	4.49×10^{27}	±0.00
	P vs. U	2.83×10^{62}	±0.00	2.01×10^{62}	±0.00	1.43×10^{62}	±0.00
	N vs. U	3.04×10^{34}	±0.00	2.21×10^{34}	±0.00	1.59×10^{34}	±0.00
normative arousal	ANOVA	3.27×10^{54}	±0.00	2.43×10^{54}	±0.00	1.80×10^{54}	±0.00
	N vs. P	5.99×10^{35}	±0.00	4.35×10^{35}	±0.00	3.12×10^{35}	±0.00
	N vs. U	1.89×10^{37}	±0.00	1.37×10^{37}	±0.00	9.79×10^{36}	±0.00
	P vs. U	0.21	±0.00	0.29	±0.04	0.37	±0.00

normative complexity	ANOVA	0.02	±0.01	0.04	±0.01	0.07	±0.01
	N vs. P	0.16	±0.07	0.21	±0.03	0.29	±0.00
	P vs. U	0.16	±0.07	0.22	±0.03	0.29	±0.00
	N vs. U	0.15	±0.07	0.21	±0.03	0.28	±0.00
apparent contrast	ANOVA	0.05	±0.01	0.10	±0.01	0.17	±0.01
	N vs. P	0.32	±0.00	0.42	±0.00	0.53	±0.00
	P vs. U	0.36	±0.00	0.48	±0.00	0.60	±0.00
	N vs. U	0.15	±0.08	0.21	±0.03	0.28	±0.00
jpeg size	ANOVA	0.03	±0.01	0.05	±0.01	0.09	±0.01
	N vs. P	0.20	±0.06	0.27	±0.03	0.35	±0.00
	P vs. U	0.19	±0.06	0.25	±0.03	0.34	±0.00
	N vs. U	0.16	±0.07	0.21	±0.03	0.28	±0.00
ratings valence	ANOVA	4.05x10^{33}	±0.00	2.96x10^{33}	±0.00	2.13x10^{33}	±0.00
	N vs. P	9.30x10^{7}	±0.00	7.39x10^{7}	±0.00	5.57x10^{7}	±0.00
	P vs. U	1.87x10^{13}	±0.00	1.37x10^{13}	±0.00	9.92x10^{12}	±0.00
	N vs. U	4.92x10^{13}	±0.00	3.61x10^{13}	±0.00	2.60x10^{13}	±0.00
ratings arousal	ANOVA	3.43x10^{8}	±0.00	3.10x10^{8}	±0.00	2.53x10^{8}	±0.00
	N vs. P	6.95x10^{7}	±0.00	5.54x10^{7}	±0.00	4.19x10^{7}	±0.00
	N vs. U	1.51x10^{9}	±0.00	1.17x10^{9}	±0.00	8.69x10^{8}	±0.00
	P vs. U	97.90	±0.00	102.00	±0.00	95.90	±0.00

Note. N: neutral; P: pleasant; U: unpleasant.

Picture ratings

At the end of the main experiment, participants were asked to provide valence and arousal ratings of all the pictures previously shown. Each image was presented for 167 ms, followed by a Self-Assessment-Manikin (Bradley & Lang, 1994) ranging from 1 (low arousal - unpleasant valence) to 9 (high arousal - pleasant valence).

Mean ratings of emotional valence and arousal collected after the main experiment are reported in *Supplementary Table S1*. A one-way rbANOVA$_{RM}$ on **valence** ratings showed a significant effect of emotion ($F_{1, 19}$ = 256, p < .001, ξ = .85, CI$_{95\%}$ [.81, 1.00]). Post-hoc tests showed higher ratings for pleasant relative to neutral (ψ = -1.19, p = 7.26x10^{-8}, g = -1.98, CI$_{95\%}$

[-2.59, -1.42]) and unpleasant scenes (ψ = 3.86, p = 1.72x10^{-10}, g = 3.53, CI$_{95\%}$ [2.60, 4.38]), as well as for neutral compared to unpleasant scenes (ψ = 2.55, p = 2.21x10^{-12}, g = 3.69, CI$_{95\%}$ [2.80, 4.66]). A Bayesian ANOVA$_{RM}$ confirmed that the data favored the model with the effect of emotion (H_1) 2.96 x 10^{33} times more than the null model (H_0) (see *Supplementary Table S2*). Bayesian paired-sample *t*-tests confirmed higher ratings for pleasant relative to neutral (BF$_{10}$ = 7.39 x 10^7 ±0.00%) and unpleasant scenes (BF$_{10}$ = 1.37 x 10^{13} ±0.00%), as well as for neutral vs. unpleasant scenes (BF$_{10}$ = 3.61 x 10^{13} ±0.00%).

A one-way rbANOVA$_{RM}$ on **arousal** ratings showed a significant effect of emotion ($F_{2,\ 29}$ = 60.40, p < .001, ξ = .89, CI$_{95\%}$ [.65, .99]). Post-hoc tests showed lower arousal for neutral compared to pleasant (ψ = -1.84, p = 9.49x10^{-9}, g = -1.96, CI$_{95\%}$ [-2.44, -1.53]) and unpleasant scenes (ψ = -2.55, p = 2.09x10^{-9}, g = -2.28, CI$_{95\%}$ [-2.99, -1.66]). Contrary to the normative ratings, arousal between pleasant and unpleasant scenes was also statistically significant, with higher arousal ratings for unpleasant compared to pleasant scenes (ψ = -0.70, p = .001, g = -0.80, CI$_{95\%}$ [-1.24, -0.33]). A Bayesian ANOVA$_{RM}$ confirmed that the data favored H_1 3.10 x 10^8 times more than H_0. Bayesian paired-sample *t*-tests confirmed lower arousal ratings for neutral relative to pleasant (BF$_{10}$ = 5.54 x 10^7 ±0.00%) and unpleasant scenes (BF$_{10}$ = 1.17 x 10^9 ±0.00%). The significant difference in arousal ratings between pleasant and unpleasant scenes was also confirmed (BF$_{10}$ = 102.00 ±0.00%).

Results of SSVEP amplitude at 6 Hz

Our *a priori* hypothesis stated that SSVEP amplitude at the stimulation frequency should not differ across *emotion* conditions, because no regularity was presented. Regarding the effect of *stimulus type*, we explored the possibility that 6 Hz amplitude could be different between original and scrambled scenes due to the processing of meaningful content in the former but not the latter picture types.

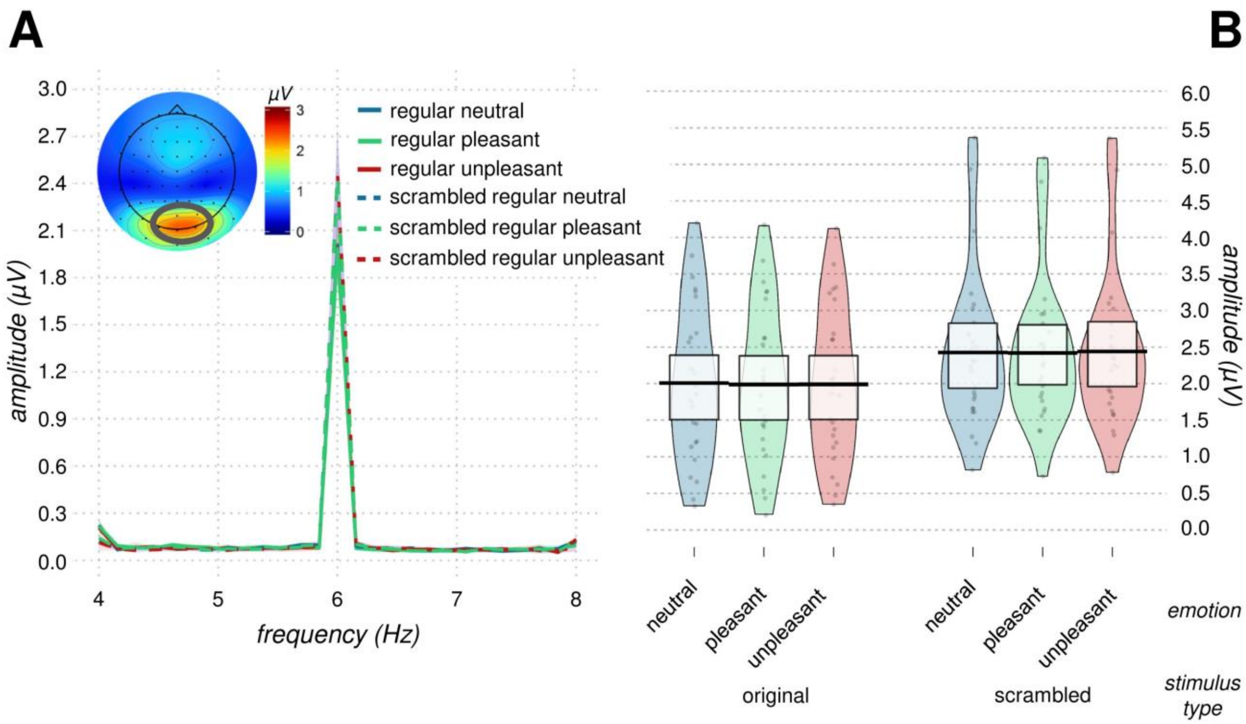

Figure S1. Results of the 6 Hz signals. (**A**) Grand average FFT-amplitude spectra derived from EEG signals at each participant's best four-electrode cluster for the 6 Hz signal for the different experimental conditions (*blue*: neutral; *green*: pleasant; *red*: unpleasant). The shaded areas around the means indicate 95% confidence intervals. The electrode cluster selected for the analysis is highlighted in the topography. (**B**) Amplitude values of the 6 Hz signals for each participant (single dots) and experimental condition. Average amplitude values are marked by horizontal black lines and 95% Bayesian highest density interval (HDI) represented as white boxes.

As shown in *Supplementary Figure S1*, 6 Hz amplitude was maximal at a cluster of central occipital electrodes (e.g., Oz, O1, O2, PO3, POz, PO4). Amplitude values are reported in

Supplementary Table S3. A 2 (*stimulus type*: concrete, scrambled) x 3 (*emotion*: neutral, unpleasant, pleasant) rbANOVA on 6 Hz amplitude values revealed no significant main effect of *stimulus type* (F = 3.08, p = .084, ξ = .218, $CI_{95\%}$ [.00, .47]) or *emotion* (F = 0.001, p = .999, ξ = .002, $CI_{95\%}$ [.00, .31]), as well as no significant *stimulus type x emotion* interaction (F = 0.007, p = .997, ξ = .218, $CI_{95\%}$ [.00, .46]). A 2 x 3 Bayesian $ANOVA_{RM}$ confirmed that the null model ought to be preferred to the model with the main effect of emotion (BF_{10} = 0.03 ±0.01%). However, it also showed that the observed SSVEP amplitudes at 6 Hz were more likely to be explained by the model with only the main effect of *stimulus type*, although anecdotally (BF_{10} = 3.00 ±0.00%) (see *Supplementary Table S4*).

Table S3. SSVEP amplitude values (in μV) at the driving frequency of 6 Hz.

condition	amplitude
irregular	1.94 (0.25)
regular neutral	1.95 (0.26)
regular pleasant	1.94 (0.26)
regular unpleasant	1.94 (0.26)
scrambled regular neutral	2.24 (0.18)
scrambled regular pleasant	2.26 (0.17)
scrambled regular unpleasant	2.26 (0.18)

Note. 20% trimmed means and standard errors (in parentheses).

Differences in stimulus processing between original and scrambled images also seem to be evident in involved brain areas, as revealed by a dissimilarity in the topographical distributions of the signal as well as differences in source estimates of the SSVEPs for the different experimental conditions. When calculating the dissimilarity between average topographies of the 6 Hz signals for the regular concrete conditions and the regular scrambled images, a dissimilarity between both topographies was revealed (GMD = 0.425). Compared to an empirical distribution of permuted GMD indices, this value is statistically significant ($p < .001$) (see *Supplementary Figure S2A*).

With respect to source estimations of the 6 Hz signal, the preprocessed EEG data were filtered with a 5.5-6.5 Hz Hamming windowed-sinc FIR bandpass filter (filter order 4096) and sources were estimated as described in the main text. Statistical analysis revealed differences in four clusters as a function of the factor *stimulus type* ($p_{voxel} < .001$, $p_{cluster} < .05$, whole-brain FWE-corrected; see *Supplementary Figure S2B*). As observable in the contrast effect estimates of the significant clusters, three clusters showed stronger source power when original images were presented: in left fusiform gyrus, left V4 (MNI coordinates of peak voxel: $x = -40$, $y = -68$, $z = -18$, $F_{peakvoxel} = 63.66$, 680 voxels in cluster), right fusiform gyrus, right V4 (MNI coordinates of peak voxel: $x = 38$, $y = -74$, $z = -12$, $F_{peakvoxel} = 51.75$, 293 voxels in cluster) and left V1, V2 (MNI coordinates of peak voxel: $x = -8$, $y = -82$, $z = 10$, $F_{peakvoxel} = 36.66$, 307 voxels in cluster). One cluster showed stronger source power when scrambled images were presented, namely right primary V1 (MNI coordinates of peak voxel: $x = 10$, $y = -94$, $z = 2$, $F_{peakvoxel} = 51.42$, 409 voxels in cluster). There were no significant clusters for the factor *emotion* or the interaction *emotion* x *stimulus* type.

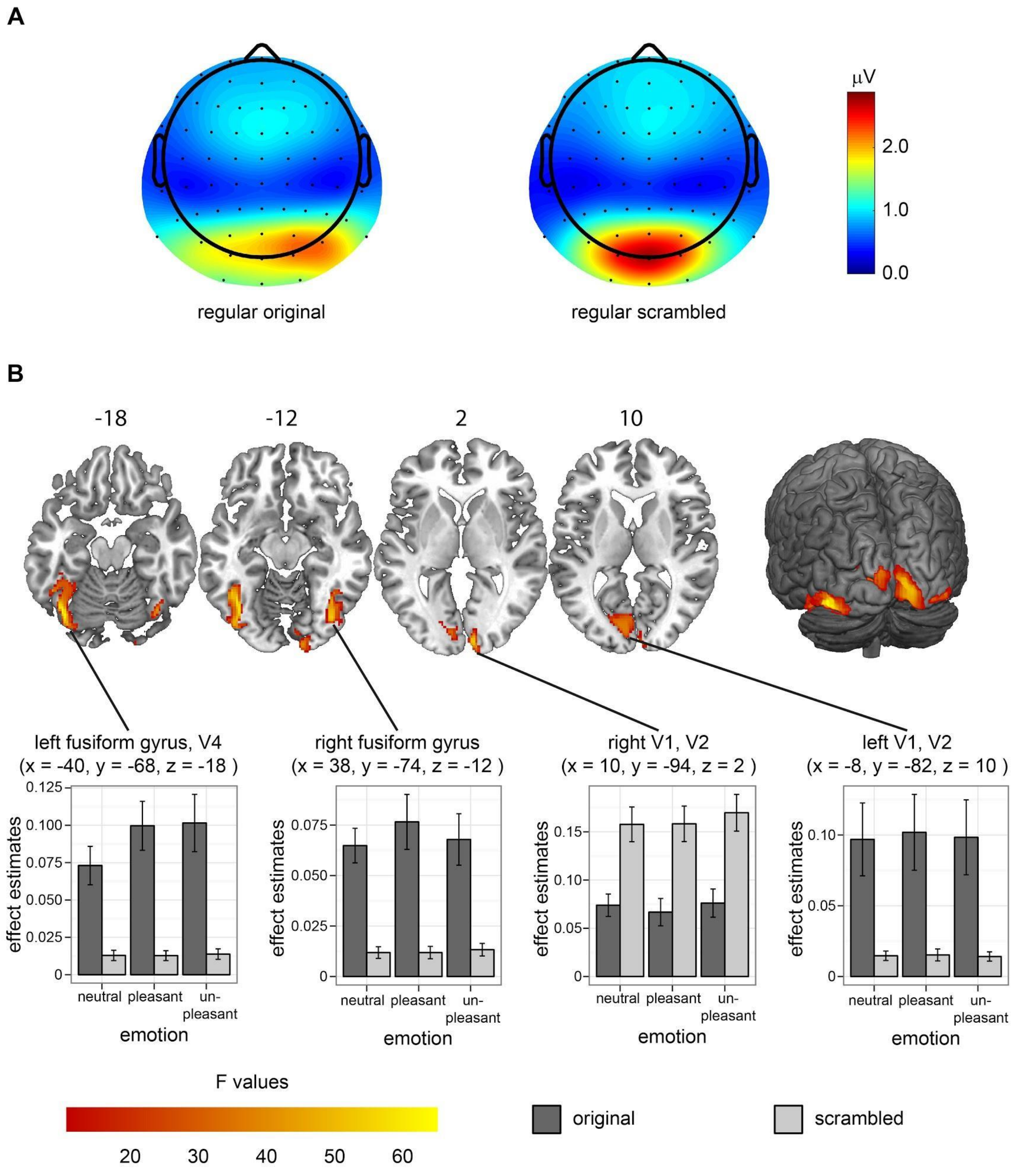

Figure S2. Results of the 6 Hz signals. (**A**) Grand Mean topographical representation of the 6 Hz signals averaged for the regularly presented original and scrambled images. (**B**) Results of the source localization. Significant source amplitude differences modulations are shown for the factor stimulus type (whole brain family-wise-error-corrected, $p_{voxel} < .001$, $p_{\text{cluster}} < .05$) in different axial slices and projected to a standard brain. Numbers above slices represent MNI coordinates of axial slice. Effect estimates are shown for each cluster in bar graphs. Error bars represent SEM.

These results confirm that the flickering presentation of complex pictures successfully elicited a steady-state response in occipital areas (Bekhtereva, Craddock, & Müller, 2015; Bekhtereva & Müller, 2015; Hindi Attar & Müller, 2012; Keil et al., 2003). Given that 6 Hz signals were comparable across emotional conditions, the observed emotion-related differences for the 2 Hz signal (see main text) must stem from specific processing differences between the regularly presented emotional images and not from differential processing of the whole picture stream itself. Moreover, these regions responded more intensively to scrambled as opposed to original scenes – as evidenced by overall larger 6 Hz amplitude values –, perhaps due to participants' effortful attempt to extract meaningful information from such perceptually noisy information. Topographies and source estimation results additionally revealed enhanced activation of visual low-level areas in response to scrambled pictures whereas concrete images activated higher order occipito-temporal areas, consistent with a large number of studies highlighting the involvement of these cortical regions during high-level scene processing (Grill-Spector, Kourtzi, & Kanwisher, 2001; Nasr et al., 2011).

Table S4. Bayes factors (BF_{10}) and percentage of proportional errors (% pe) for each model of interest, obtained by using JZS priors with different scaling factors.

model	*r= 1*		*r= 0.707*		*r= 0.5*	
	BF_{10}	% pe	BF_{10}	% pe	BF_{10}	% pe
stimtype	2.30	±0.00	3.00	±0.00	3.73	±0.00
emo	0.02	±0.01	0.03	±0.01	0.06	±0.01

stimtype x emo	0.001	±0.70	0.006	±0.85	0.03	±0.91

Note. *stimtype*: stimulus type; *emo*: emotion.

Post-hoc analyses

Unbalanced presence of humans across emotion categories

During a previous submission, two reviewers raised the possibility that other, non-controlled stimulus features in our stimulus set might have caused the observed increased 2 Hz amplitude in response to pleasant pictures. Specifically, they referred to a *"lack of control of semantic coherency"* (Reviewer #1) which could have resulted, for instance, in pleasant scenes containing *"more (or more visible) humans"* (Reviewer #2). We counted the stimuli displaying living objects according to whether they showed human bodies vs. not: 28/9 for neutral pictures, 42/1 for pleasant pictures, and 36/5 for unpleasant pictures. Thus, most of our pre-selected stimuli displayed human bodies.

A chi-square test of independence showed a statistically significant difference across emotion categories ($\chi^2_{2,\ N\ =\ 121}$ = 8.86, p = .012, Cramer's V = .27, $CI_{95\%}$ [0.01, 0.43]), anecdotally supported by the Gûnel and Dickey BF^{I}_{10} = 2.34 ±0.00%. Paired comparisons showed human/non-human unbalance between pleasant and neutral pictures ($\chi^2_{1,\ N\ =\ 80}$ = 8.80, p = .003, Cramer's V = .33, $CI_{95\%}$ [0.12, 0.51]; BF^{I}_{10} = 14.91 ±0.00%), whereas the differences between neutral vs. unpleasant ($\chi^2_{1,\ N\ =\ 78}$ = 1.94, p = .163, Cramer's V = .16, $CI_{95\%}$ [0.00, 0.37]; BF^{I}_{10} = 0.54 ±0.00%) or pleasant and unpleasant scenes ($\chi^2_{1,\ N\ =\ 84}$ = 3.08, p = .079, Cramer's V = .19 $CI_{95\%}$ [0.00, 0.39]; BF^{I}_{10} = 0.58 ±0.00%) were not significant.

These results allow us to neither confirm nor disconfirm the possibility that the presence of humans in our pre-selected stimuli might have contributed to the selective 2 Hz amplitude increase in response to pleasant scenes. Future studies specifically addressing this issue would be informative.

Preferential processing of nude bodies

One of the reviewers additionally suggested that *"pleasant scenes contained images of sexual intercourse presumably showing nude bodies"* (Reviewer #2), pointed to previous research reporting amplitude enhancement of early ERP components to this class of stimuli (Alho, Salminen, Sams, Hietanen, & Nummenmaa, 2015; Hietanen, Kirjavainen, & Nummenmaa, 2014; Hietanen & Nummenmaa, 2011), and wondered whether our 2 Hz results would reflect a similar phenomenon.

We believe this alternative explanation to be highly unlikely in our study, given that nude bodies were displayed in only 14 out of 43 pleasant pictures with living objects. In any case, even if 2 Hz amplitude increases in response to pleasant pictures were due to quick attention allocation towards nude bodies, this would still be compatible with our hypothesis that pleasant emotional information (here conveyed by stimuli eliciting sexual arousal) is preferentially selected and quickly processed in situations of high perceptual load. Please note that a general arousal account cannot easily explain the observed pattern of results, mainly because our participants rated *unpleasant* scenes as more arousing compared to pleasant ones (see *Supplementary Table S1*).

General remarks

While we appreciate the importance of carefully controlling the stimulus set with respect to as many dimensions as possible, we believe that categorizations based on semantic content are difficult to operationalize because rather subjective. For example, which dimensions should be used for semantic categorization? Is the presence of living/non-living objects sufficient or should a distinction between humans and non-humans be made (Crouzet, Joubert, Thorpe, & Fabre-Thorpe, 2012)? Is it important whether the displayed actions are happening indoor or outdoor

(Rousselet, Joubert, & Fabre-Thorpe, 2005)? Furthermore, these categories are typically used in isolation but, in real-world examples, they overlap (Banno & Saiki, 2015): for instance, erotic pictures are mostly set indoor, therefore one may wonder whether they are preferentially processed because there are human bodies or because they are set indoor. More complex scene categorization taxonomies (Xiao et al., 2013) may be suitable for machine learning algorithms but much less for the emotional distinction investigated here.

References

Alho, J., Salminen, N., Sams, M., Hietanen, J. K., & Nummenmaa, L. (2015). Facilitated early cortical processing of nude human bodies. *Biological Psychology*, *109*, 103–110.

Banno, H., & Saiki, J. (2015). The Processing Speed of Scene Categorization at Multiple Levels of Description: The Superordinate Advantage Revisited. *Perception*, *44*(3), 269–288.

Bates, E., D'Amico, S., Jacobsen, T., Székely, A., Andonova, E., Devescovi, A., … Tzeng, O. (2003). Timed picture naming in seven languages. *Psychonomic Bulletin & Review*, *10*(2), 344–80.

Bekhtereva, V., Craddock, M., & Müller, M. M. (2015). Attentional bias to affective faces and complex IAPS images in early visual cortex follows emotional cue extraction. *NeuroImage*, *112*, 254–266.

Bekhtereva, V., & Müller, M. M. (2015). Affective facilitation of early visual cortex during rapid picture presentation at 6 and 15 Hz. *Social Cognitive and Affective Neuroscience*, *10*(12), 1623–33.

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.

Bradley, M. M., & Lang, P. J. (1994). Measuring Emotion: The Self-Assessment Mannequin and the Semantic Differential. *Journal of Behavior Therapy and Experimental Psychiatry*, *25*(1), 49–59.

Crouzet, S. M., Joubert, O. R., Thorpe, S. J., & Fabre-Thorpe, M. (2012). Animal Detection Precedes Access to Scene Category. *PLOS ONE*, *7*(12), e51471.

Field, A. P., & Wilcox, R. R. (2017). Robust statistical methods: A primer for clinical psychology and experimental psychopathology researchers. *Behaviour Research and Therapy*, *98*, 19–38.

Grill-Spector, K., Kourtzi, Z., & Kanwisher, N. (2001). The lateral occipital complex and its role in object recognition. *Vision Research*, *41*(10–11), 1409–1422.

Gûnel, E., & Dickey, J. (1974). Bayes factors for independence in contingency tables. *Biometrika*, *61*(3), 545–557.

Hietanen, J. K., Kirjavainen, I., & Nummenmaa, L. (2014). Additive effects of affective arousal and top-down attention on the event-related brain responses to human bodies. *Biological Psychology*, *103*, 167–175.

Hietanen, J. K., & Nummenmaa, L. (2011). The Naked Truth: The Face and Body Sensitive N170 Response Is Enhanced for Nude Bodies. *PLOS ONE*, *6*(11), e24408.

Hindi Attar, C., & Müller, M. M. (2012). Selective attention to task-irrelevant emotional distractors is unaffected by the perceptual load associated with a foreground task. *PloS One*, *7*(5), e37186.

Jamil, T., Ly, A., Morey, R. D., Love, J., Marsman, M., & Wagenmakers, E.-J. (2016). Default “Gunel and Dickey” Bayes factors for contingency tables. *Behavior Research Methods*, 1–15.

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.

Keil, A., Gruber, T., Müller, M. M., Moratti, S., Stolarova, M., Bradley, M. M., & Lang, P. J. (2003). Early modulation of visual perception by emotional arousal: Evidence from steady-state visual evoked brain potentials. *Cognitive, Affective & Behavioral Neuroscience*, *3*(3), 195–206.

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (2008). *International affective picture system (IAPS): Affective ratings of pictures and instruction manual. Technical Report A-8*. University of Florida, Gainesville, FL.

Marin, M. M., & Leder, H. (2013). Examining complexity across domains: relating subjective and objective measures of affective environmental scenes, paintings and music. *PloS One*, *8*(8), e72412.

Nasr, S., Liu, N., Devaney, K. J., Yue, X., Rajimehr, R., Ungerleider, L. G., & Tootell, R. B. H. (2011). Scene-selective cortical regions in human and nonhuman primates. *The Journal of Neuroscience*, *31*(39), 13771–13785.

New, J., Cosmides, L., & Tooby, J. (2007). Category-specific attention for animals reflects ancestral priorities, not expertise. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(42), 16598–603.

Proverbio, A. M., Del Zotto, M., & Zani, A. (2007). The emergence of semantic categorization in early visual processing: ERP indices of animal vs. artifact recognition. *BMC Neuroscience*, *8*(1), 24.

Rousselet, G. A., Joubert, O. R., & Fabre-Thorpe, M. (2005). How long to get to the “gist” of real-world natural scenes? *Visual Cognition*, *12*(6), 852–877.

Schettino, A., Keil, A., Porcu, E., & Müller, M. M. (2016). Shedding light on emotional perception: Interaction of brightness and semantic content in extrastriate visual cortex. *NeuroImage*, *133*, 341–353.

Thorpe, S., Fize, D., & Marlot, C. (1996). Speed of processing in the human visual system. *Nature*, *381*(6582), 520–522.

Wessa, M., Kanske, P., Neumeister, P., Bode, K., Heissler, J., & Schönfelder, S. (2010). EmoPics: Subjektive und psychophysiologische Evaluationen neuen Bildmaterials für die klinisch-biopsychologische Forschung. *Zeitschrift Für Klinische Psychologie Und Psychotherapie*, *39*(1/11), 77.

Xiao, J., Hays, J., Russell, B. C., Patterson, G., Ehinger, K., Torralba, A., & Oliva, A. (2013). Basic level scene understanding: categories, attributes and structures. *Frontiers in Psychology*, *4*.

Appendix

IAPS image codes. *Neutral*: 1101, 1112, 1303, 1505, 1645, 1675, 1935, 1945, 2032, 2309, 2372, 2487, 2521, 2616, 2695, 2704, 2752, 3005.2, 3310, 3550.2, 4130, 4274, 4275, 4325, 4561, 5395, 5455, 5661, 5970, 6000, 6837, 6930, 7092, 7247, 7484, 7503, 7600, 7620, 7830, 8010, 8121, 8325, 9150, 9422, 9635.2, 9913. *Unpleasant*: 1050, 1300, 1931, 1932, 2730, 3019, 3103, 3150, 3195, 3212, 3213, 3250, 3500, 6021, 6212, 6312, 6315, 6530, 6540, 6550, 6560, 6834, 8480, 9250, 9321, 9414, 9600, 9620, 9622, 9623, 9902, 9904, 9921. *Pleasant*: 1650, 4220, 4290, 4311, 4490, 4643, 4652, 4664, 4668, 4693, 4695, 4698, 4810, 5470, 5621, 5626, 5629, 7405, 7650, 8030, 8034, 8080, 8158, 8161, 8163, 8170, 8178, 8179, 8180, 8185, 8186, 8191, 8193, 8200, 8206, 8251, 8300, 8341, 8370, 8400, 8470, 8490, 8492, 8499, 8501.

EmoPics image codes: *Neutral*: 136, 151, 205, 316. *Unpleasant*: 209, 210, 211, 212, 213, 227, 229, 230, 231, 235, 238, 244, 245, 247, 249, 252, 326. *Pleasant*: 52, 53, 61, 63, 64.