



UvA-DARE (Digital Academic Repository)

The time course of natural scene perception with reduced attention

Groen, I.I.A.; Ghebreab, S.; Lamme, V.A.F.; Scholte, H.S.

DOI

[10.1152/jn.00896.2015](https://doi.org/10.1152/jn.00896.2015)

Publication date

2016

Document Version

Final published version

Published in

Journal of Neurophysiology

License

Article 25fa Dutch Copyright Act (<https://www.openaccess.nl/en/policies/open-access-in-dutch-copyright-law-taverne-amendment>)

[Link to publication](#)

Citation for published version (APA):

Groen, I. I. A., Ghebreab, S., Lamme, V. A. F., & Scholte, H. S. (2016). The time course of natural scene perception with reduced attention. *Journal of Neurophysiology*, 115(2), 931-946. <https://doi.org/10.1152/jn.00896.2015>

General rights

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

The time course of natural scene perception with reduced attention

 Iris I. A. Groen,¹ Sennay Ghebreab,^{1,2} Victor A. F. Lamme,¹ and H. Steven Scholte¹

¹Amsterdam Brain and Cognition Center, Department of Brain and Cognition, University of Amsterdam, Amsterdam, The Netherlands; ²Intelligent Systems Lab Amsterdam, Institute of Informatics, University of Amsterdam, Amsterdam, The Netherlands

Submitted 21 September 2015; accepted in final form 9 November 2015

Groen II, Ghebreab S, Lamme VA, Scholte HS. The time course of natural scene perception with reduced attention. *J Neurophysiol* 115: 931–946, 2016. First published November 25, 2015; doi:10.1152/jn.00896.2015.—Attention is thought to impose an informational bottleneck on vision by selecting particular information from visual scenes for enhanced processing. Behavioral evidence suggests, however, that some scene information is extracted even when attention is directed elsewhere. Here, we investigated the neural correlates of this ability by examining how attention affects electrophysiological markers of scene perception. In two electro-encephalography (EEG) experiments, human subjects categorized real-world scenes as man-made or natural (full attention condition) or performed tasks on unrelated stimuli in the center or periphery of the scenes (reduced attention conditions). Scene processing was examined in two ways: traditional trial averaging was used to assess the presence of a categorical manmade/natural distinction in event-related potentials, whereas single-trial analyses assessed whether EEG activity was modulated by scene statistics that are diagnostic of naturalness of individual scenes. The results indicated that evoked activity up to 250 ms was unaffected by reduced attention, showing intact categorical differences between manmade and natural scenes and strong modulations of single-trial activity by scene statistics in all conditions. Thus initial processing of both categorical and individual scene information remained intact with reduced attention. Importantly, however, attention did have profound effects on later evoked activity; full attention on the scene resulted in prolonged manmade/natural differences, increased neural sensitivity to scene statistics, and enhanced scene memory. These results show that initial processing of real-world scene information is intact with diminished attention but that the depth of processing of this information does depend on attention.

attention; natural scenes; electro-encephalography/event-related potential; image statistics

THE HUMAN VISUAL SYSTEM TRANSFORMS low-level retinal input into high-level neural representations using several parallel processing streams (Kravitz et al. 2013; Ungerleider and Mishkin 1982). Despite its impressive processing power, however, the brain is capacity limited and cannot represent all information from our rich and complex environment at once (Marois and Ivanoff 2005). The selection of relevant information and suppression of irrelevant information are regulated by visual attention, which can modulate neural activity at multiple levels of visual processing (Kastner and Pisk 2004).

Evidence suggests that the need for attentional selection increases when the information that is processed becomes increasingly complex. For instance, attention is necessary to construct multifeature representations for solving complex visual search tasks (Wolfe 1998) but not for single feature

representations. The powerful phenomenon of “change blindness,” whereby large changes in real-world environments go unnoticed when attention is directed elsewhere (Rensink 2001), suggests that, in highly complex scenes, attending to a specific scene element severely limits visual processing of the other elements in that scene.

However, research on rapid perception of real-world scenes shows that some information, such as the presence of a target object (Potter 1975) or scene category (Greene and Oliva 2009), can be extracted with minimal effort. The processing of this information appears to be unaffected by diminished attention; object detection remains intact when subjects have to process two scenes in parallel instead of a single scene, effectively doubling the visual input (Rousselet et al. 2002). Similarly, performance is intact when complex scenes are presented in a dual-task paradigm, even though discrimination of more simple stimuli suffers under these circumstances (Li et al. 2002). Thus, paradoxically, even for rich and complex natural scenes, some information is able to escape the attentional bottleneck (Braun 2003).

What is the neural basis for the resilience of complex scene processing to diminished attention? It has been proposed that attention can operate in multiple cognitive subsystems, such as sensory processing, working memory, or response selection, depending on the nature of the stimuli and the task at hand (Luck et al. 2000). Attentional selection may only be necessary in a given subsystem if that subsystem suffers from interference from other stimuli or tasks (Lavie 1995). Event-related potentials (ERPs) are particularly useful to elucidate at which stage of processing attentional selection takes place because their high temporal resolution can identify the time window at which stimulus processing is affected even if there is no response requirement to that stimulus (Luck 2005). Sensory-level interference effects, for example, as observed in spatial cuing experiments, can manifest as early as the C1 or P1 (between 60 and 100 ms after stimulus onset), whereas effects related to, for instance, working memory updating (reflected in the P300) or the resolution of response conflict take place after this early period in the “post-perceptual” processing stage (Luck et al. 2000).

Although the early and late effects of attention across various tasks have been well characterized for simple, artificial stimuli, much less is known about how attention impacts the neural processing of real-world scenes (Peelen and Kastner 2014). Here, we explored the time course of attentional interference on scene perception by examining the effects of reduced attention on activity evoked by real-world scenes. In two experiments, we compared the level of scene processing during an explicit categorization task with tasks performed on simul-

Address for reprint requests and other correspondence: I. Groen, National Institute of Mental Health, 10 Center Dr., Bethesda, MD 20892 (e-mail: iris.groen@nih.gov).

taneously presented, unrelated stimuli, thereby reducing the amount of attention directed at the scenes. In Experiment 1, subjects performed a central letter task that was matched in difficulty to the scene categorization task, as well as a more difficult letter task that taxed working memory. In Experiment 2, we examined the spatial extent of the attention effects by introducing a condition in which attention was diverted to a peripherally presented stimulus. Critically, the same scenes were shown in each condition, such that the influence of attention could be isolated. If, despite their complexity, sensory analysis of scenes can indeed occur unperturbed even in the presence of interfering stimuli and competing tasks, the degree of scene processing as reflected in early ERP time windows should remain intact with diminished attention.

The degree of scene processing was evaluated by means of two types of ERP analysis. First, we performed traditional ERP trial averaging to test whether a categorical difference between manmade and natural scenes was present in evoked activity in each condition. Second, single-trial analysis was used to test to what extent attention affected neural sensitivity to image-specific scene statistics. Many scene properties are correlated with low-level statistical regularities, for example in the spatial frequency spectrum (Torralba and Oliva 2003) or the distribution of local contrast (Brady and Field 2000). In a previous study, we examined how these statistics affected manmade/natural categorization on a large set of diverse real-world scenes (Groen et al. 2013). We found that the mean local contrast energy of a scene predicted response times, whereas a parameter describing the shape of the local contrast distribution, which we dubbed “spatial coherence,” correlated most strongly with naturalness categorization. Importantly, these statistics also parametrically modulated sensory ERP amplitudes, and spatial coherence in particular predicted how well a category could be decoded from single-trial ERP responses to the scenes (Groen et al. 2013). Thus local contrast statistics provide a useful single-image quantification of scene naturalness that can be used to probe more fine-grained neural processing of scene information. By comparing ERP sensitivity to these statistics across the different conditions, we sought to provide convergent evidence for neural scene processing under different attentional manipulations.

METHODS

Experiment 1

Subjects. Nineteen participants (4 men, aged 19–27 yr; mean 21.5, SD = 2.4) took part in Experiment 1. To ensure that participants had normal or corrected-to-normal vision, a Landolt C visus test was administered before the experiment; a visual acuity of 0.8 in both eyes was required for participation. The experiment was approved by the Ethical Committee of the University of Amsterdam, and all participants gave written, informed consent before participation and were compensated for their time with research credits or money (7 euro/h). After preprocessing, one participant was excluded from analysis because s/he failed to respond on >25% of the trials. Two more participants were excluded based on low electroencephalography (EEG) quality (excessive movement, visible as artifacts in >10 min of raw data, or a large amount of drift likely attributable to a malfunctioning reference).

Experimental procedure. EEG was recorded while subjects were presented with stimulus displays that consisted of manmade and natural scenes and letters. A continuous stream of black letters was

displayed in the center of the screen on a gray background. On average, every nine letters, a white letter was presented instead of a black letter, and a scene was simultaneously presented behind it (Fig. 1A; note that, for illustrative purposes, the letters are displayed much larger in the figure than in the actual experiment; see *Trial design* for details). Upon this event, subjects were required to perform a two-alternative forced choice (2-AFC) task based on one of three different task instructions: 1) indicate whether the scene is manmade or natural (“scene task”); 2) indicate whether the white letter is a vowel or consonant (“vowel task”); or 3) indicate whether the white letter was the same as two white letters before (“2-back task”; Fig. 1B). Tasks were blocked, and each block contained 300 trials. Each subject completed two blocks per task, first each of the three tasks once and then again repeated in the same order. Task order was counterbalanced across subjects.

Subjects were instructed to respond to white letters/scene onsets only and to attend to the black letters to ensure fixation. They indicated their responses with the index fingers of each hand using custom-made button boxes taped to the chair armrests. Response mappings were counterbalanced across the task repetitions, i.e., within subjects. Each block lasted 9.7 min and was divided in mini-blocks of 166 s. Mini-blocks started with a 2,000-ms fixation cross. After each mini-block, a feedback screen was presented displaying the number of correct responses for each choice (e.g., “you responded correctly for 95% of the vowels”) and how many times they failed to respond before the onset of the next trial. The feedback screen was colored according to subjects’ performance on the preceding mini-block: green if >90% of responses was correct, orange if it was 80–90%, and red if it was <80%. Subjects initiated the onset of a new mini-block with a button press.

Memory posttest. To evaluate whether the task manipulation was successful in reducing attention to the scenes, subjects performed an additional experiment after removal of the EEG cap and electrodes, which assessed how well subjects remembered scenes they observed during the main EEG experiment. Subjects were presented with scenes they had seen previously during the EEG experiment as well as new scenes (see *Scene selection*). They indicated whether they recognized the scene by pressing the button box with the left or right index finger. Scenes were displayed one by one on the same monitor using the same display settings as in the EEG experiment, except that the letter was now omitted. Also, each scene was visible until the subjects responded to allow the subjects enough time to recognize the scene. However, subjects were instructed to respond based on their first impression and to not linger on the images. The test took about 5 min on average. Importantly, it was not announced to the participant until the EEG experiment was completed.

Computation of image statistics. For each scene, we calculated two image statistics based on distributions of local contrast (Scholte et al. 2009): contrast energy (CE) and spatial coherence (SC). These statistics are derived in a biologically realistic way by summing simulated output of contrast-sensitive receptive fields. The details of the model are described in detail in various other places (Ghebreab et al. 2009; Groen et al. 2013; Scholte et al. 2009). Briefly, the model performs the following calculations: 1) conversion of the RGB layers of the image to color-opponent space (grayscale, red-green, blue-yellow; Koenderink et al. 1972); 2) local contrast detection within each color-opponent space using multiscale exponential contrast filters (Zhu and Mumford 1997) whose outputs are rectified and divisively normalized using the lateral geniculate nucleus suppressive field approach (Bonin et al. 2005); 3) construction of one contrast-intensity map for each opponent space by selecting one specific filter output per image location using minimal reliable scale selection (Elder and Zucker 1998; Ghebreab et al. 2009); 4) pooling of the contrast responses across a selection of the image (the central 1.5° for CE, the central 5° for SC), taking the average response for CE, and the coefficient of variation (means \pm SD) for SC; 5) averaging of the CE and SC values

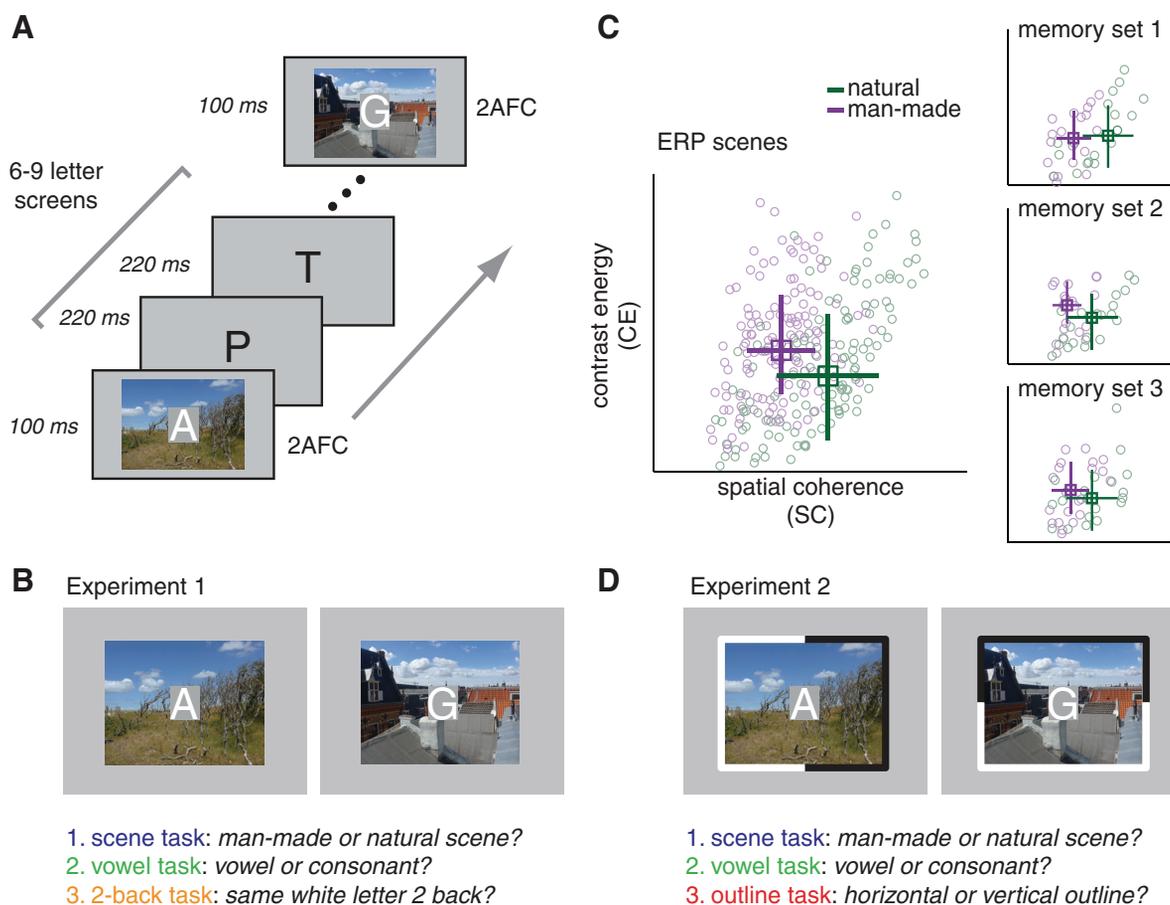


Fig. 1. Experimental design. *A*: trial design of Experiment 1. On each trial, a white letter and a scene replaced a continuous stream of black letters, and a 2-alternative forced choice (2AFC) was required. Note that, for illustrative purpose, letters and scenes are displayed considerably larger than in the actual experiment. The example pictures shown here are from the author's personal collection and were not used in the actual experiments. *B*: in Experiment 1, subjects performed a scene task, vowel task, and an 2-back task. *C*: contrast energy (CE) and spatial coherence (SC) values for each scene image (scattered circles) shown in the electro-encephalography (EEG) experiments. *Left*: event-related potential (ERP) scenes that were used in all tasks and analyzed in the ERP analyses. *Right*: 3 matched sets of scenes that were each uniquely presented in 1 task and tested again in the memory posttest. Squares = medians, lines = standard deviations within each global category (manmade and natural). *D*: in Experiment 2, an outline appeared along with the scene and white letter, and a 2AFC outline discrimination task replaced the 2-back letter task.

across color-opponent spaces to obtain one CE and SC value per scene.

These computational steps have been chosen to maximize the correlation of the CE and SC with two parameters derived from a Weibull function fitted to the distribution of contrast values (Scholte et al. 2009). CE is a biologically realistic approximation of the distribution mean (the scale parameter of the function), whereas SC is an approximation of its shape (the degree to which the function describes a power law or a Gaussian distribution). These two statistics thus capture information about the overall presence of edges in a scene (CE) and higher-order correlations between them (SC). As a result, images with low CE/SC values contain strong figure-ground segregation (i.e., because of the presence of a coherent object with sharp edges), whereas images with high values are cluttered or textured. The correlation of CE and SC with scene naturalness thus stems from the fact that natural scenes tend to contain clutter or otherwise unorganized structure, whereas urbanized, manmade environments are more orderly (Groen et al. 2013).

Scene selection. Scenes were selected from a larger set of 1,600 scenes used by Groen et al. (2013). In that study, subjects ($n = 14$) categorized scenes as either manmade or natural, and their average response served as the ground-truth labels in the present study. For the present experiment, we randomly selected 500 scenes (250 manmade, 250 natural) that were consistently rated across all previous observers.

These 500 scenes were divided as follows: 120 scenes (60 manmade, 60 natural) were set apart and not presented in the EEG experiment but only in the memory test, thus serving as new scenes. The other 380 scenes were used in the EEG experiment. Of these, 130 manmade and 130 natural scenes were designated "ERP scenes" and were presented once in each task block in the EEG experiment. Only these scenes were used for the ERP analyses, which are thus based on the exact same scenes in all conditions. The remaining 120 scenes were designated as "memory scenes" and were split in 3 sets of 40 scenes (20 manmade, 20 natural), which were presented only within one specific task (once per block). These scenes were excluded from ERP analysis but were presented again in the memory posttest, randomly intermixed with the new scenes. By presenting these scenes uniquely in one condition, we could determine to what extent our task manipulations affected subsequent scene memory. Memory sets were counterbalanced across tasks, such that there was no systematic relation between memory set and task condition.

In addition, scene selection was constrained such that average CE and SC values within the manmade and natural categories did not differ between ERP scenes, memory scenes, or new scenes (see Fig. 1C; for Kruskal-Wallis tests of equal medians and Bartlett's tests of equal variances, all $P > 0.20$).

Trial design. Stimuli were presented using Presentation software (<http://www.neurobs.com>). To construct a trial, nine black letters were

drawn randomly without replacement from a predetermined set and arranged in a mini letter train. Then the fourth-sixth letters of this train were randomly picked and replaced with a white letter. White letters were never directly repeated, and the first black letter of a train was never the same as the last black letter of the previous train. Black and white letters were presented for 216 ms. Additionally, at the onset of the white letter, a natural scene was presented for a duration of 100 ms. Scenes were 640×480 pixels and were presented on the center of a 19-inch ASUS monitor with a frame rate of 60 Hz and a screen resolution of 1920×1080 pixels. Subjects were seated at a distance of 90 cm from the screen, such that the scenes subtended $\sim 9 \times 7^\circ$ of visual angle. The letters were displayed on top of a small gray box of $\sim 55 \times 35$ pixels (the size of the box varied slightly depending on the letter shown), subtending $\sim 0.8 \times 0.5^\circ$ of visual angle.

For the scene and vowel tasks, the white letters were four vowels (A, E, O, and U) and four consonants (S, T, G, and P). White letters were drawn randomly on each trial, resulting in on average 50% vowels and 50% consonants being presented. For the black letters, 12 other consonants were used (C, D, H, J, K, L, N, Q, X, Y, and Z). For the 2-back task, we created two sets of eight white letters (A, F, G, P, S, R, U, W and B, D, E, L, M, O, T, V). Every six trials, letters were drawn from one of the two white letter sets to create a 2-back sequence in which a repetition occurred on either the fourth, fifth, or sixth position. Thus in total 50 target trials ("same white letter as 2 white letters back") were presented. The two white letter sets were alternated to assure that a target trial could not be confused with a letter of previous trials. The black letters consisted of the remaining letters in the alphabet (C, D, H, I, J, K, N, Q, X, Y, and Z).

Behavioral data analysis. Responses and reaction times were recorded. For behavioral analysis of the EEG experiment, both ERP scenes and memory scenes were included. Trials for which subjects either failed to respond or responded before 200 ms or after 1,200 ms were excluded from analysis (mean = 1.5%, median = 0.8%, min = 0.2%, max = 6% of all 1,800 trials). For each subject and each task block, sensitivity (d') was computed based on response accuracy per scene category. Hits were designated as correct categorization of a manmade scene (scene task), a vowel (vowel task), or a two back repeated letter (2-back task). False alarms were designated as incorrect categorization of a manmade scene as natural, a vowel as consonant, or a nonrepetition as two back, respectively. To avoid infinite z -values for 100% hits and 0% false alarms, the scores were corrected by setting these proportions to 99% and 1% (Macmillan and Creelman 1991). Reaction times were averaged across trials to compute mean and median RT across subjects for each task block.

For analysis of the memory posttest, d' was calculated by designating correctly recognized scenes as hits and designating incorrectly recognized new scenes as false alarms. The data were analyzed in MatLab (Mathworks, Natick, MA) and SPSS 17.0 (IBM, Armonk, NY).

EEG recording and preprocessing. EEG was recorded with a 64-channel Active Two EEG system (Biosemi Instrumentation, Amsterdam, The Netherlands, www.biosemi.com) at a sample rate of 2,014 Hz. The EEG setup was similar to that of our previous studies (Groen et al. 2012a, 2012b, 2013). We used caps with an extended 10–20 layout modified with 2 additional occipital electrodes (I1 and I2, which replaced F5 and F6). Eye movements were recorded with additional electrooculograms (EOG).

Preprocessing was done in Brain Vision Analyzer 2 (BVA2) and included the following steps: 1) low-pass filtering at 115.2 Hz and 24 dB/octave to prevent aliasing due to down sampling; 2) down sampling of the raw data to 256 Hz and offline referencing to the average of two external electrodes placed on the mastoids; 3) a high-pass filter at 0.1 Hz (12 dB/octave), a low-pass filter at 30 Hz (24 dB/octave), and two notch filters at 50 and 60 Hz; 4) automatic removal of deflections > 300 ms (after visual inspection, this threshold was raised for some subjects with very high ERP amplitudes); 5) segmentation into epochs from -100 to 500 ms from scene onset; 6) ocular

correction based on the EOG activity (Gratton 1983); 7) baseline correction between -100 and 0 ms; 8) automated artifact rejection (maximal voltage $50 \mu\text{V}$, minimal/maximal amplitudes $-75/75 \mu\text{V}$, lowest activity $0.50 \mu\text{V}$); and 9), conversion of the obtained ERPs to current source density responses (Perrin 1989). For the temporal filtering, infinite impulse response filters were used, which are implemented as phase shift-free Butterworth filters in BVA2. Including trials that were rejected based on behavior, median rejection rate was 87 out of 1,560 ERP scene trials (min = 11 trials, max = 345 trials); in total, 6% of the data was rejected. After preprocessing, the single-trial ERP amplitude values were imported into MatLab (Mathworks) for statistical analysis.

ERP analysis. To subtract out effects of response preparation (and because the behavioral data did not indicate any difference between repetitions; see RESULTS) trials were averaged across the two blocks of data obtained for each task. To examine whether information about scene naturalness was present in the evoked activity in each of the different tasks, we computed grand average ERPs and difference waves based on the ERP scenes and compared these across tasks (memory scenes were excluded from ERP analysis). First, average ERP amplitude for manmade and for natural scenes was computed for each subject and each task separately. The natural ERP was then subtracted from the manmade ERP to obtain a difference wave for each subject and task. Per electrode and time point, average difference wave amplitude was tested against zero with a two-tailed, one-sample t -test. P values were corrected for multiple comparisons with an FDR correction at $\alpha = 0.05$. In addition, we report only significant differences lasting for at least six consecutive samples (~ 23 ms) (Johnson and Olshausen 2003; Rieger et al. 2013). To compare the difference waves across tasks, two-tailed paired t -tests were conducted for each pair of tasks, and the same FDR threshold was applied as for the difference wave analysis. Although all analyses were conducted without preselection of electrodes, i.e., for all 64 electrodes simultaneously, we focused our analyses on the posterior (occipital and parietal) electrodes, which overlay the cortical areas known to be involved in visual processing and categorization.

Single-trial ERP regression. To investigate how image statistics affected evoked activity to the scenes, we ran regression analysis of single-trial ERP amplitude on CE and SC values of the scenes (Groen et al. 2013). Regression analyses were performed separately for each experiment, subject, electrode, and time point, with ERP amplitude values as the dependent variable and CE and SC values as independent variables. The regression was performed simultaneously for all trials but with separate predictor columns for each task. The regression model thus contained 3×260 rows (780 trials) and 7 predictor columns (constant, CE scene task, SC scene task, CE vowel task, SC vowel task, CE 2-back task, SC 2-back task). Predictor columns were z -scored independently. Missing trials were marked as NaN and thereby ignored in the regression analysis.

Each regression resulted in a single measure of overall explained variance (r^2), as well as regression weights (β -coefficients) specific to each task and image statistic. For each electrode and time point, the significance of the β -coefficients was determined by testing whether the mean β -coefficient across subjects was larger than zero using two-tailed one-sample t -tests. Resulting P values were FDR corrected for multiple comparisons (electrodes, time points, and tasks) at $\alpha = 0.05$. For consistency with the difference wave analysis discussed above, we only report β -coefficients that were significant for at least six consecutive samples. Comparisons in regression coefficients between tasks were carried out by means of paired t -tests using the same FDR threshold as for the individual β -coefficients.

Experiment 2

Subjects. Eighteen participants (4 men, aged 18–26 yr; mean 21.4, SD = 2.5) took part in the second experiment. Inclusion criteria and compensation were as in Experiment 1. Two participants were ex-

cluded from analysis based on EEG quality (excessive drift). The experiment was approved by the Ethical Committee of the University of Amsterdam, and all participants gave written, informed consent before participation.

Experimental procedure, trial design, and scene selection. To examine whether the late effects of attention on ERP amplitude (see RESULTS, *Experiment 1*) were due to peripheral suppression or to explicit scene categorization, we ran a second experiment in which we manipulated peripheral rather than focal attention. Procedures were identical to Experiment 1, except that, apart from letters and scenes, a black-and-white frame surrounding the scene was presented (Fig. 1D). The letter 2-back task was replaced by a peripheral discrimination task in which subjects indicated whether the outline was oriented vertically or horizontally (“outline task”). Outlines were drawn by presenting a black and a white box behind the scene that extended it by 10 pixels ($\sim 0.1^\circ$) on each side. Outline orientation was randomly determined on each trial. In addition, it was randomly determined which half of the outline was black, with the constraint that a particular outline configuration (e.g., “horizontal orientation with white top and black bottom”) was never directly repeated. In all tasks, letters were presented according to the settings of the vowel task in Experiment 1, except that, not only the scenes, but also the white letters and the outlines were now all presented for 100 ms to ensure that stimulation duration was equal across tasks. Visual stimulation was thus identical across the three tasks, and the only difference was task requirement. The same scenes were used as in Experiment 1, and a memory posttest was again administered.

Data analysis. Behavioral data and ERP analysis was identical to the first experiment. For the behavioral analysis of the EEG experiment, again a small number of trials were excluded (mean = 0.5%, median = 0.4%, min = 0%, max = 1.8% of all 1,800 trials). In the outline task, correct categorization of a horizontal outline was designated as a hit and categorization of a horizontal outline as vertical as a false alarm. For the ERP analysis, median trial rejection rate was 88 out of 1,560 trials (min = 5 trials, max = 256 trials); in total, 7% of the data was rejected. To compare difference wave onsets of Experiment 2 with Experiment 1, two-tailed independent-sample *t*-tests were conducted between the average manmade/natural difference waves from each experiment. The resulting *P* values were FDR corrected for multiple comparisons across time samples and electrodes at $\alpha = 0.05$.

RESULTS

Experiment 1

Rationale. All subjects performed a manmade/natural scene categorization task (full attention on scene) and two letter categorization tasks, while presented with the same set of scene stimuli (Fig. 1, A and B). We used two different letter tasks to divert attention from the scenes, an easy vowel categorization task and a more challenging 2-back task. To assess the extent to which attention was successfully manipulated in each condition, we separately tested how well subjects remembered (Wolfe et al. 2007) intermixed scenes that were unique to each task (Fig. 1C) in an unannounced old-new posttest (see METHODS).

Task and memory performance. Task performance during EEG recording was analyzed using a two-way repeated-measures ANOVA with the factor task (scene, vowel, or 2-back) and repetition (first or second block). We observed a main effect of task on d' [$F(2,14) = 33.0, P < 0.001$]. There was no effect of repetition [$F(1,15) = 0.03, P = 0.85$] or interaction between task and repetition [$F(2,15) = 0.02, P = 0.96$]. Pair-wise *t*-tests on mean d' across repetitions indicated that

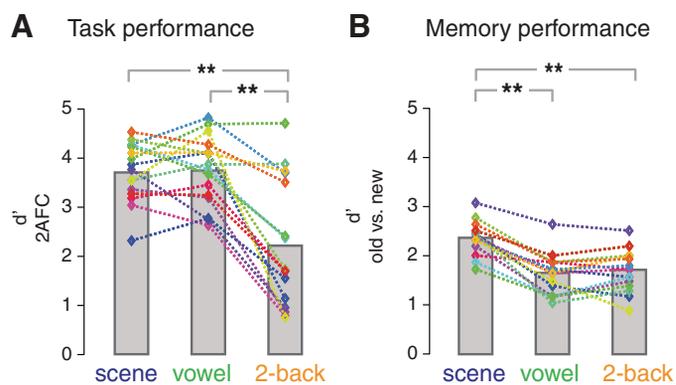


Fig. 2. Behavioral results of Experiment 1. A: performance on the scene, vowel, and 2-back tasks during EEG recording, including mean d' across 2 repetitions for each subject (colored lines) and the mean value across subjects (gray bars). B: memory posttest performance, including mean d' for each subject (lines) and mean across subjects (gray bars) for images shown during either the scene, vowel, or 2-back task. $**P < 0.001$.

subjects performed less well in the 2-back task compared with the other two tasks [2-back vs. scene, $t(15) = 6.8, P < 0.001$; 2-back vs. vowel, $t(15) = 6.0, P < 0.001$; scene vs. vowel, $t(15) = 0.25, P = 0.81$; Fig. 2A]. There were no significant main effects on RT [all $F(2,15) < 2.5$, all $P > 0.10$; mean RT \pm SD scene: 525 ± 44 ms; vowel: 531 ± 37 ms; 2-back: 502 ± 58 ms; median RT \pm SD scene: 500 ± 40 ms; vowel: 510 ± 33 ms; 2-back: 477 ± 62 ms]. These results show that, although response speed was similar across tasks, the 2-back task was more difficult than the vowel task, as intended.

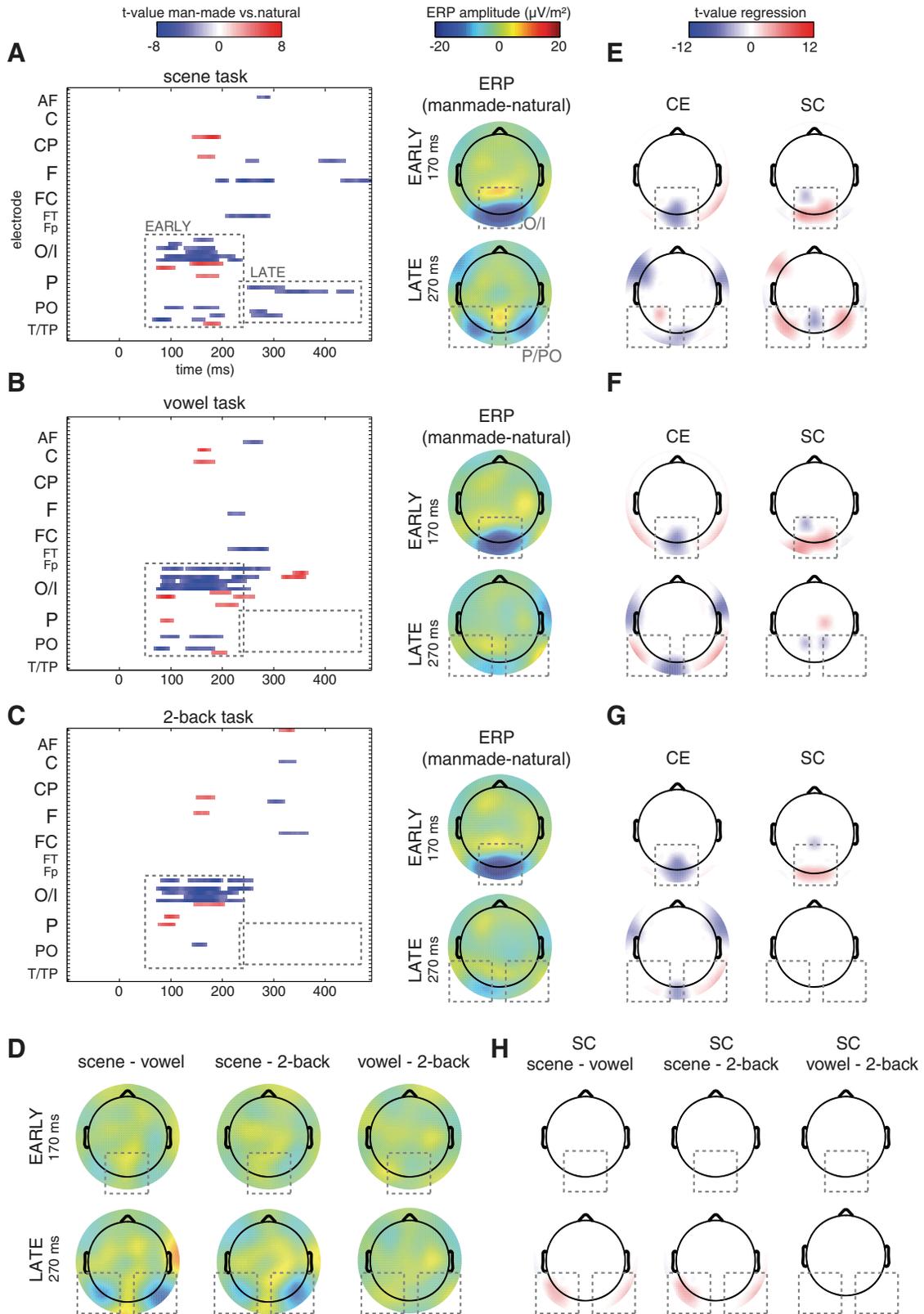
For the memory posttest performance, we observed a main effect of task [1-way repeated-measures ANOVA on d' : $F(2,14) = 52.3, P < 0.001$]. Subjects more accurately recognized scenes they had seen in the scene task compared with the other tasks [scene vs. vowel, $t(15) = 9.7, P < 0.001$; scene vs. 2-back, $t(15) = 7.1, P = 0.001$], which did not differ from one another [vowel vs. 2-back, $t(15) = -0.9, P = 0.34$; Fig. 2B]. Importantly, there was not a particular set of memory scenes that was inherently better remembered than other sets [repeated-measures ANOVA on memory sets: $F(1,15) = 0.5, P = 0.58$]. These results show that scenes attended during categorization were indeed better remembered. However, although the 2-back task was experienced as more difficult, there was no difference in scene memory between the 2-back and vowel task, suggesting that attention was diverted from the scene in the two letter tasks to a similar degree.

ERP results. To examine when scene information emerged in evoked activity in each condition, difference waves between manmade and natural scenes were computed for each electrode separately and statistically evaluated using one-sample *t*-tests. In addition, we extracted scene-specific single-trial ERP amplitudes to examine neural sensitivity to scene statistics. We first present the whole-scalp results and then discuss individual electrodes that most clearly demonstrate the observed effects.

Whole-scalp difference waves. In the scene task, the first significant deflection of the difference wave between manmade and natural scenes appeared at 70 ms after stimulus onset at occipital midline electrode POz (Fig. 3A). This difference was negative, indicating a larger ERP for natural compared with manmade scenes, and was part of a large cluster of occipital electrodes (Oz, O1, O2, I1, I2, and PO4) that showed significant negative deflections up to 200–250 ms. This negative

deflection was accompanied by positive deflections at several parietal electrodes (Pz, P1, P3, and P10). Together, these differences reflected a strong source of activity at the back of the scalp that was maximal between 150 and 200 ms. We define this as the early effect. Around 250 ms, we observed a

second, late effect of scene naturalness centered on lateral parietal-occipital and parietal electrodes (PO7-PO8 between 260–324 ms; P3–P7 between 258–464 ms). This effect was present as two lateral deflections from ~250 ms onward (Fig. 3A).



If we compare these results to the letter tasks (vowel and 2-back), we see that the early effect was also present there. The deflections at the occipital cluster were nearly identical, starting at 74 ms at POz (vowel task) and at 78 ms at Oz (2-back task). For the vowel task (Fig. 3*B*), the early cluster consisted of the same electrodes as in the scene task (with the exception of I2), and the early differences were again significant up to 250–300 ms. For the 2-back task, the early differences were slightly more focal (restricted to I1, I2, O1, O2, and Oz), with a brief deflection at PO4 (Fig. 3*C*); they also lasted up to 250 ms, and the topographic scalp distributions were again visually similar.

Notably, however, in these reduced attention conditions, the late effect disappeared completely; there were no significant deflections at the parietal-occipital and parietal channels from 250 ms onward (compare the gray dashed boxes labeled “LATE” in Fig. 3, *A–C*), and the lateral sources observed in the scene task were no longer visible on the scalp (compare boxes labeled “P/PO” in the topographic plots of Fig. 3, *A–C*).

To demonstrate the task-induced differences more clearly, direct comparisons between the manmade/natural difference waves for each task are shown in Fig. 3*D*. At the early interval, there were no significant differences between tasks at early occipital electrodes. At the late interval, the difference waves were significantly more negative in the scene task compared with the vowel and 2-back tasks at occipital electrode O2 and at perioccipital electrodes (scene vs. vowel: PO8, P5, and P10; scene vs. 2-back: PO4, PO8, P6, and P7). There were no significant differences at any of the occipital or parietal electrodes between the vowel and 2-back tasks.

These results show that a difference between manmade and natural scenes is consistently present in early evoked activity in all three tasks, even when attention was diverted to the letters. However, late and more widespread activity is selectively present in the scene task only. We next examined whether global scene information was also present at the single-trial level, measured as sensitivity to CE and SC, which respectively summarize the presence of local contrasts in an image and higher-order correlations between them (see METHODS).

Whole-scalp sensitivity to scene statistics. For the single-trial regression analysis, all trials were modeled simultaneously, but tasks were represented in different independent variables such that task-specific regression weights for CE and SC were obtained (see METHODS). To examine whether the regression model described the data well, we first consider the model fits (explained variance) in each subject. Effect sizes were greatest at electrode Oz: average r^2 at this electrode reached a maximum of 19% at 105 ms after stimulus onset (range across subjects 8–47%, 90–117 ms). At this electrode

and time point, model fits were significant in every single subject [mean $F(6, 1,553) = 60.0$, all $P < 1e-17$]. We next considered how the two statistics separately influenced ERP amplitude in each task by examining the regression weights for CE and SC.

In the scene task, significant β -coefficients for both CE and SC were found at occipital electrodes in the early time interval (Fig. 3*E*) (first significant CE coefficient at 70 ms at POz; at 170 ms, significant CE coefficients at I2, Oz, O1, O2, and P10; first significant SC coefficient at 66 ms at POz; at 170 ms, significant SC coefficients at I2, Oz, O1, O2, P1, and PO4). At the later time interval, the β -coefficients for CE still centered on occipital sites (at 270 ms, significant coefficients at I2, I1, I2, Oz, O2, P3, and P2), whereas those for SC were also found on lateral electrodes (at 270 ms, significant coefficients at PO7, PO8, P6, P7, P8, POz, and Pz). In the vowel (Fig. 3*F*) and 2-back (Fig. 3*G*) tasks, the early occipital effects were identical to the scene task for both CE and SC. The later effects for SC, however, were no longer present. This task-related difference is clearly revealed when the β -coefficients are compared across tasks (Fig. 3*H*). At the two time points representing the early and late effects, no significant difference was found between the tasks in CE coefficients, but SC coefficients were present only in the scene task at lateral channels P7, PO7, and P10.

These results show that, at the single-trial level, early sensitivity to scene statistics is maintained with diminished attention but that late sensitivity at higher-level channels is modulated by attention. These temporally distinct effects of scene information on neural activity were also separated across the scalp; the automatic, early effect was centered on central visual electrodes, whereas the late, attention-dependent effect extended toward lateral occipital-parietal sites. We next illustrate these effects at representative electrodes.

Individual electrodes. Figure 4 shows the results for the early, midline occipital electrode Oz. At Oz, manmade and natural ERPs in the scene task differed significantly from ~80 ms onward (Fig. 4*A*). The difference waves for the three tasks at this electrode overlapped completely, and the onsets of significance of the difference wave were identical between the tasks (Fig. 4*B*). Interestingly, regression coefficients for CE became significant exactly at these time points (see vertical gray dashed line in Fig. 4*A–C*), whereas SC started to affect ERP amplitude slightly later in time (Fig. 4*C*). Scene information was thus reliably present at this early visual channel regardless of task, at both the grand average and single-trial ERP level.

Two late lateral parietal-occipital electrodes (PO7 and PO8) are illustrated in Fig. 5. At both PO7 and PO8, the manmade/natural difference in the scene task appeared at ~150 ms,

Fig. 3. Whole-scalp ERP and regression results for Experiment 1. *A*, left: t values of the difference wave between manmade and natural scenes at each electrode (rows) and time sample (columns) in the scene task. Positive t values indicate that the ERP for manmade scenes is larger than for natural scenes (positive difference wave); negative t values reflect a negative difference wave. The gray dashed boxes illustrate the early and late effects. Electrode abbreviations: AF, anterior-frontal; C, central; CP, central-parietal; F, frontal; FC, frontal-central; FT, frontal-temporal; Fp, frontal pole; O/I, occipital; P, parietal; PO, parietal-occipital; T, temporal; TP, temporal-parietal. Right: whole-scalp difference wave amplitude representative of the early and late effects in *A*. The gray dashed boxes illustrate the early occipital and the late parietal-occipital electrode locations. *B* and *C*: same as *A*, but for the vowel and 2-back task, respectively. *D*: difference in difference wave (manmade minus natural) amplitude between tasks (scene minus vowel; scene minus 2-back; vowel minus 2-back) at the early and late time points. Color scale is the same as in *A–C*, right column. *E*: early and late whole-scalp maps of significant β -coefficients for CE and SC, as expressed in their t values. Positive values indicate a positive correlation of CE or SC with ERP amplitude; negative values are a negative correlation. T values are only displayed on the scalp if they survive FDR correction. *F* and *G*: same as *E*, but for the vowel and 2-back tasks, respectively. Note that the regression weights for SC, but not CE, are sensitive to attention (gray dashed boxes). *H*: difference in SC regression coefficients between tasks at the early and late time point. There were no significant differences between tasks for CE at these time points. Color scales and thresholds are the same as in *E–G*.

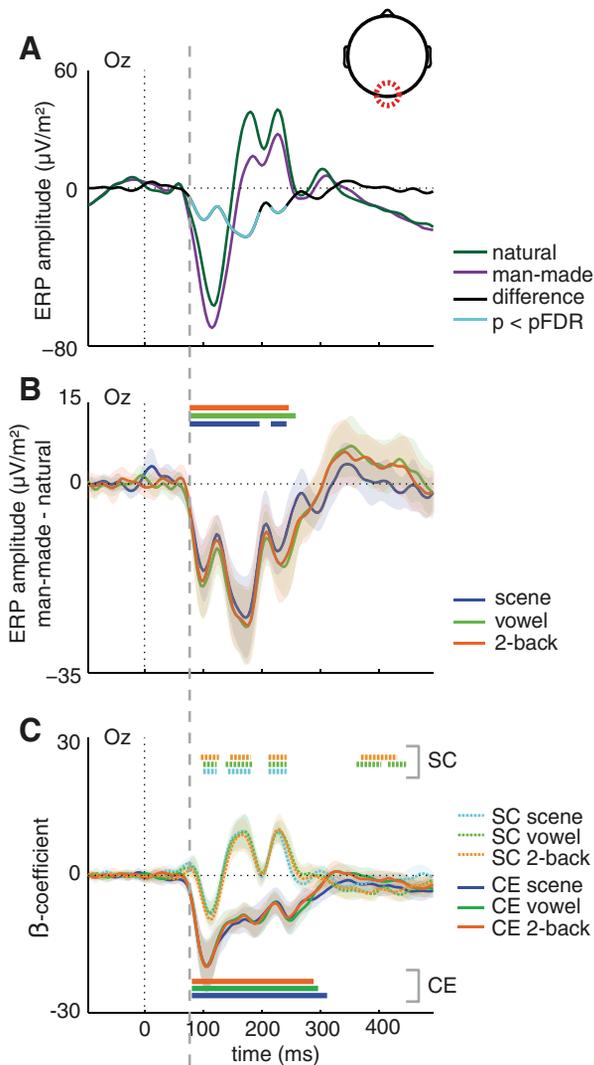


Fig. 4. Single-electrode ERP and regression results for Experiment 1 at occipital channel Oz. *A*: grand average ERPs for manmade and natural scenes and the difference wave for the scene task. Difference wave significance was determined by 1-sample *t*-tests at each time sample and electrode, controlling for multiple comparisons with FDR correction (FDR-corrected *P* value threshold = 0.0027). *B*: difference waves for all 3 tasks at Oz. Shadings indicate confidence intervals derived from the *t*-tests. Thick straight colored lines indicate significant time points. *C*: β -coefficients for CE and SC in each task. Lines indicate significant time points, controlled for multiple comparisons using FDR correction (FDR-corrected *P* value threshold = 0.0028). The gray vertical dashed line is superimposed to guide the eye to compare the onsets of the effects across the 3 panels.

reaching significance at ~ 250 ms at PO7 (Fig. 5A) and ~ 170 ms at PO8 (Fig. 5B). Although the difference waves for the three tasks were also initially overlapping, they were only significant in the scene task at the late interval (Fig. 5, C–D). In accordance, regression coefficients for CE and SC were weaker at these electrodes compared with Oz (Fig. 5, E–F). The largest coefficients beyond 170 ms were found for SC, which extended to ~ 300 – 350 ms in the scene task only.

We thus find two neural correlates of scene perception, which are spatio-temporally distinct. Initially, grand average manmade vs. natural differences and single-trial sensitivity to both CE and SC are present exclusively on occipital electrodes near the midline, whereas later in time these effects become

much more lateralized and selective to the global image property SC. Importantly, attentional manipulation selectively affects the late, lateralized neural activity, whereas the early occipital effects are automatically evoked upon the presentation of the scene.

Late lateral effects: spatial or scene-based attention? The results of Experiment 1 provide neurophysiological evidence that the brain can process scene information with diminished attention. In addition, we observed a selective enhancement of scene processing on more lateral electrodes with increased attention. Does this lateral activity reflect scene-specific processing induced by attending to the scene content? This is not unlikely, but we cannot conclude this with certainty from the results of Experiment 1 alone. The two letter tasks both required categorization of a central stimulus, whereas the scene extended more peripherally. Thus the increased lateral activity in the letter conditions may simply reflect enhanced processing of peripheral parts of the stimulus display, not necessarily enhanced scene processing. Another caveat of Experiment 1 is that the central letter stimuli were presented almost twice as long as the scenes, thereby possibly shifting the attentional time course of the letters compared with the scenes.

In Experiment 2, we aimed to further investigate the origin of the late lateral activity by including a condition in which subjects directed spatial attention more peripherally but still performed a task unrelated to scene category. We changed the visual stimulation by presenting a scene outline (Fig. 5D) along with the scene and letter and replaced the 2-back condition with an outline discrimination task. In addition, we equalized the presentation duration of the letter, scene, and outline. If the lack of lateral activity in the letter tasks was driven by pure peripheral suppression, late lateral activity should be present in the outline task. In contrast, if this activity reflects selective processing of scene information, it should only be present in the scene task.

Experiment 2

Behavior. Task performance (d') during EEG recording was analyzed using a two-way repeated-measures ANOVA with factors task (scene, vowel, or outline) and repetition (first or second block). We again observed a main effect of task [$F(2,14) = 10.7$, $P = 0.001$], and there was again no effect of repetition [$F(1,15) = 0.5$, $P = 0.50$] or interaction between task and repetition [$F(2,15) = 1.7$, $P = 0.21$]. Pair-wise *t*-tests on mean d' indicated that subjects had lower sensitivity in the outline task compared with the other two tasks [outline vs. scene, $t(15) = 5.4$, $P < 0.001$; outline vs. vowel, $t(15) = 3.2$, $P = 0.006$; scene vs. vowel, $t(15) = 0.34$, $P = 0.74$.; Fig. 5A]. This time, task also affected reaction times [mean RT: $F(2,15) = 4.7$, $P = 0.03$; median RT: $F(2,15) = 5.9$, $P = 0.02$]; subjects responded slightly slower in the outline condition compared with the other conditions (mean RT \pm SD scene: 550 ± 105 ms; vowel: 551 ± 123 ms; outline: 582 ± 77 ms; median RT \pm SD scene: 530 ± 51 ms; vowel = 532 ± 47 ms; outline = 564 ± 49 ms). This likely reflects the fact that, in the intertrial intervals, the central black letter stream kept spatial attention at fixation, whereas, upon trial onset, attention had to be relocated to the outline when the target stimulus appeared. There was no effect of repetition or an interaction between task and repetition on RT [all $F(2,15) < 1.3$, all $P > 0.28$].

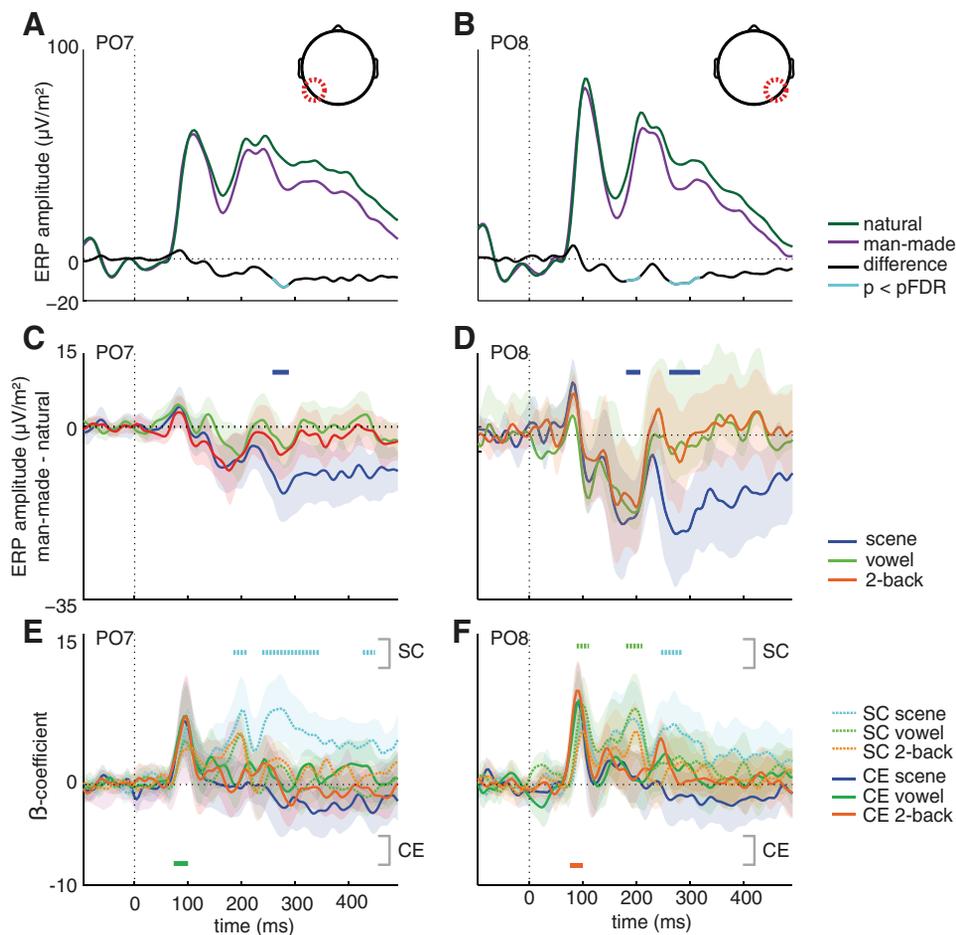


Fig. 5. Single-electrode ERP and regression results for Experiment 1, same as in Fig. 4 but for parietal-occipital electrodes PO7 (A, C, and E) and PO8 (B, D, and F).

For memory posttest performance, we again observed a main effect of task [1-way repeated-measures ANOVA on d' : $F(2,14) = 22.5$, $P < 0.001$]. Subjects again more accurately recognized scenes viewed in the scene task compared with the other two tasks [scene vs. vowel, $t(15) = 5.2$, $P < 0.001$; scene vs. outline, $t(15) = 5.5$, $P < 0.001$; vowel vs. outline, $t(15) = 1.16$, $P = 0.26$; Fig. 6B]. There was again no effect of memory set [F(1,15) = 0.7, $P = 0.52$].

These results show that, when subjects directed spatial attention peripherally in the outline task, they still remembered the scenes less well compared with the scene task. In fact, memory performance was equally impaired as in the focal

attention condition. It thus appears that the observed differences in memory performance are not driven by peripheral suppression per se but instead by attending to the scene itself, which subsequently resulted in better scene memory.

Whole scalp ERP results. The ERP averaging analysis for the scene (Fig. 7A) and vowel (Fig. 7B) tasks revealed a similar pattern of results as in Experiment 1. For both tasks, there was again an early, occipital cluster that showed a negatively deflecting difference wave (significant at POz, Oz, O1, O2, PO3, and PO4). A later, lateral deflection was again present in the scene task but not in the vowel task. However, this time, late deflections reached significance in the right hemisphere only (PO8, P6, P8, and P10).

In the new outline condition (Fig. 7C), the early negative occipital differences were again present (significant at POz, O1, O2, Oz, and PO3), and the scalp distribution of this early effect was again very similar to the other conditions. Interestingly, some significant late effects of naturalness were found in this condition, restricted to left lateral electrodes (PO3, P3, P5, P7, and P9). However, these effects were reversed in sign (manmade > natural) and lasted shorter than the previously observed lateral clusters in the scene task; the positive left lateral sources in the outline task were maximal around 245 ms. There were no significant deflections at the electrode PO8, which showed the strongest effects in the scene task.

In Fig. 7D, we again show the subtractions between the difference waves for each task. Whereas none of the electrodes showed an early difference between tasks, the late interval

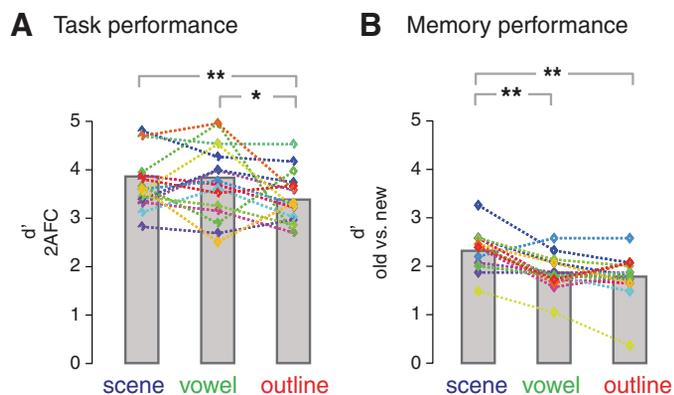


Fig. 6. Behavioral results of Experiment 2. A and B: as in Fig. 2, with the outline task replacing the 2-back task. ** $P < 0.001$. * $P < 0.01$.

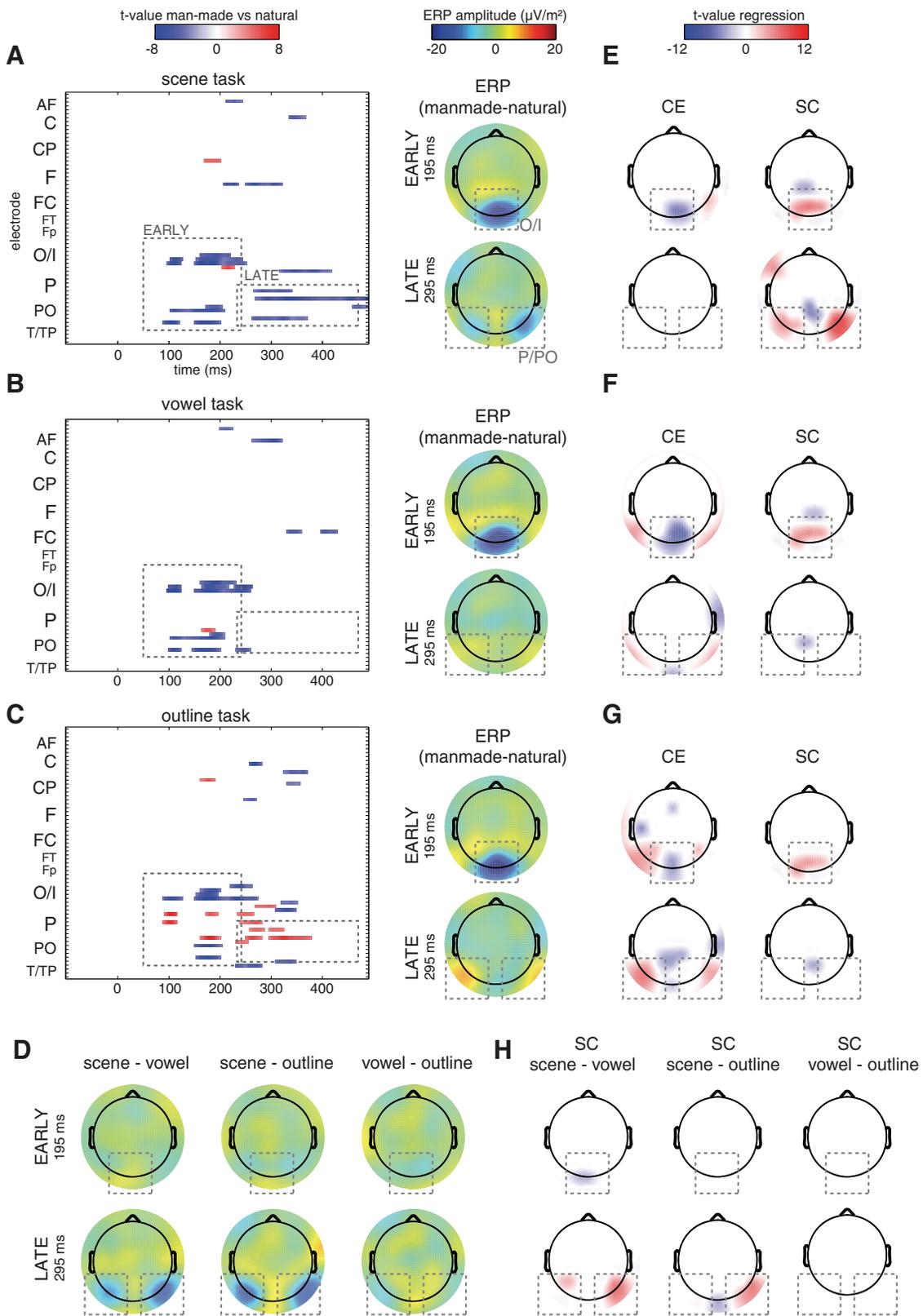


Fig. 7. Whole-scalp ERP and regression results for Experiment 2. A–H: as in Fig. 3, except that the 2-back task was replaced by the outline task. Dashed boxes are as in Experiment 1. Note that the time labels are different here because of an ~25-ms shift in overall ERP onsets between experiments (see RESULTS).

difference waves were significantly larger in the scene task compared with the vowel task at right lateral channels PO8, P6, and P8 and left lateral channel P5. Compared with the outline task, the difference waves differed significantly on both right lateral (PO8, P6, P8, P10, TP8) and left lateral channels (PO3, PO7, P5, P7, P9) although it should be noted that the latter effect could be partly explained by the reversed difference wave in the outline task.

Regression weights. In Experiment 2, maximal explained variance for the regression analysis was 23% (range 11–41% at 105 ms, range across subjects 90–117 ms), and model fits were again significant in every single subject. In the scene task, early significant β -coefficients for CE and SC were again found at occipital electrodes (Fig. 7E). As before, late coefficients were found for SC; however, similar to the difference waves, they were more right lateralized than in Experiment 1. In the vowel task (Fig. 7F), the early occipital coefficients were very similar to the scene task for both CE and SC. The later coefficients for SC, however, were again eliminated. For the outline task (Fig. 7G), the early coefficients are similar to the other tasks. However, in the later time interval, we observed a left lateralized effect of CE only. Comparison of the SC coefficients between tasks (Fig. 7H) shows that the results from Experiment 1 were replicated; SC coefficients were significant in the late interval on lateral electrodes.

Individual electrodes. At the single electrode level, we can readily see that the effects at channel Oz were replicated in Experiment 2. The early occipital difference wave in the scene task (Fig. 8A) again overlapped completely with difference waves of the other two tasks (Fig. 8B) and was accompanied by sensitivity to CE and SC (Fig. 8C). For the higher-level electrodes, the lateralization of the manmade/natural sensitivity observed at the whole scalp is visible as a difference between PO7 (Fig. 9A) and PO8 (Fig. 9B). At the right electrode, the difference between manmade and natural scenes is visible in the scene condition only, but this effect was diminished in the left electrode (Fig. 9, C and D). For the difference waves at both electrodes, the scene task time course was again mirrored in the selective neural sensitivity to SC (Fig. 9, E and F). Remarkably, at PO7 there was also sustained sensitivity to CE, in the outline task only.

Together, these ERP results replicate the early effects of Experiment 1; we observe intact early processing of scene information under attentional requirements. Regarding the late activity on the lateral electrodes, however, the results are mixed, depending on hemisphere. At right lateral electrodes, which showed the strongest manmade/natural differential activity in Experiment 1, the late effects were again only present in the scene task, and this effect was again accompanied by increased sensitivity to SC. This is consistent with the interpretation that the late lateral activity reflects scene-specific processing. However, in the left hemisphere, we observed a late effect of scene naturalness in the outline task only, which was reversed in sign compared with the right hemisphere. Importantly, this effect was accompanied by enhanced sensitivity to the contrast energy of the scenes. This result thus suggests that peripheral attention can lead to enhanced processing of some scene information, even if scene categorization is not required for the current task.

Difference wave onset comparison between experiments. We observed one other, unexpected difference in Experiment 2,

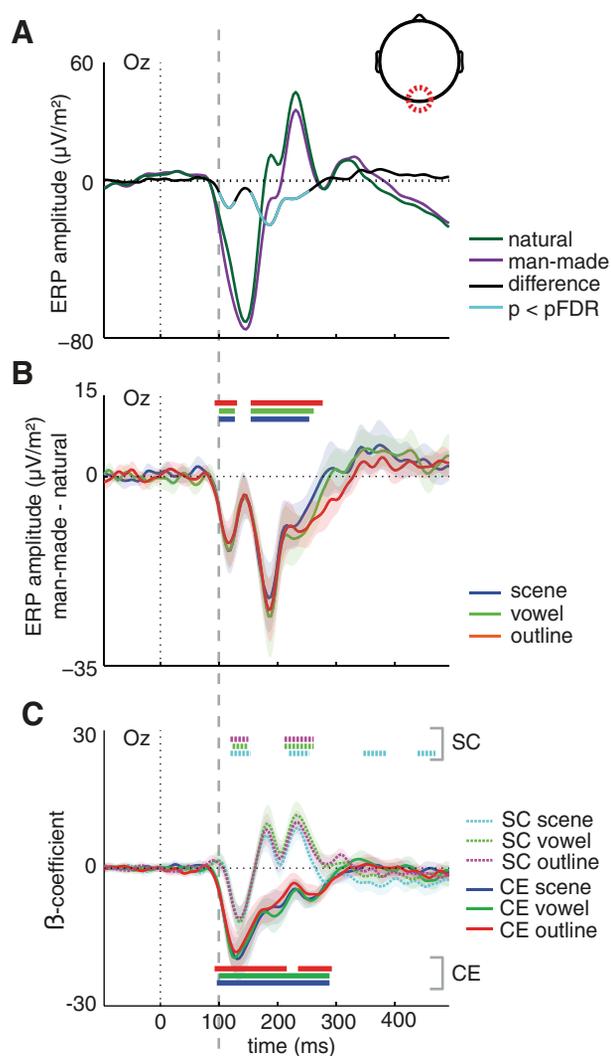


Fig. 8. Single-electrode ERP and regression results for Experiment 2 at channel Oz. A–C: same as Fig. 5, with the outline task replacing the 2-back task. For the ERP analysis, FDR-corrected P value threshold = 0.0022. For the regression analysis, FDR-corrected P value threshold = 0.0036.

namely that, in all tasks, the difference wave onsets occurred ~ 20 – 25 ms later compared with Experiment 1 (scene: 95 ms at POz; vowel: 100 ms at Oz; outline: 95 ms at Oz). Statistical comparison between experiments suggested that the entire ERP had shifted (significant difference between the mean manmade/natural difference waves across tasks on electrodes Iz, I1, O1, Oz, and P10 between ~ 70 – 100 ms). This shift was accompanied by an overall delay in RT in Experiment 2. For the scene and vowel task, which were identical between experiments, RTs were higher in Experiment 2; a two-way ANOVA with factors task (vowel/scene) and experiment (Experiment 1/Experiment 2) demonstrated a significant main effect of experiment on median RT [$F(1,31) = 5.2$, $P = 0.03$; for mean RT, the effect was not significant: $F(1,31) = 3.5$, $P = 0.07$], with no effect of task [both $F(1,31) < 0.34$, both $P > 0.56$]. Thus for both tasks there was a modest increase in RTs across experiments of about 20–25 ms (mean RT Experiment 1: 530 ± 40 ms; Experiment 2: 551 ± 55 ms; median RT Experiment 1: 506 ± 37 ms; Experiment 2: 531 ± 50 ms).

ERP onsets can be affected by participant age and by screen luminance (Bieniek et al. 2013), but subject age did not differ

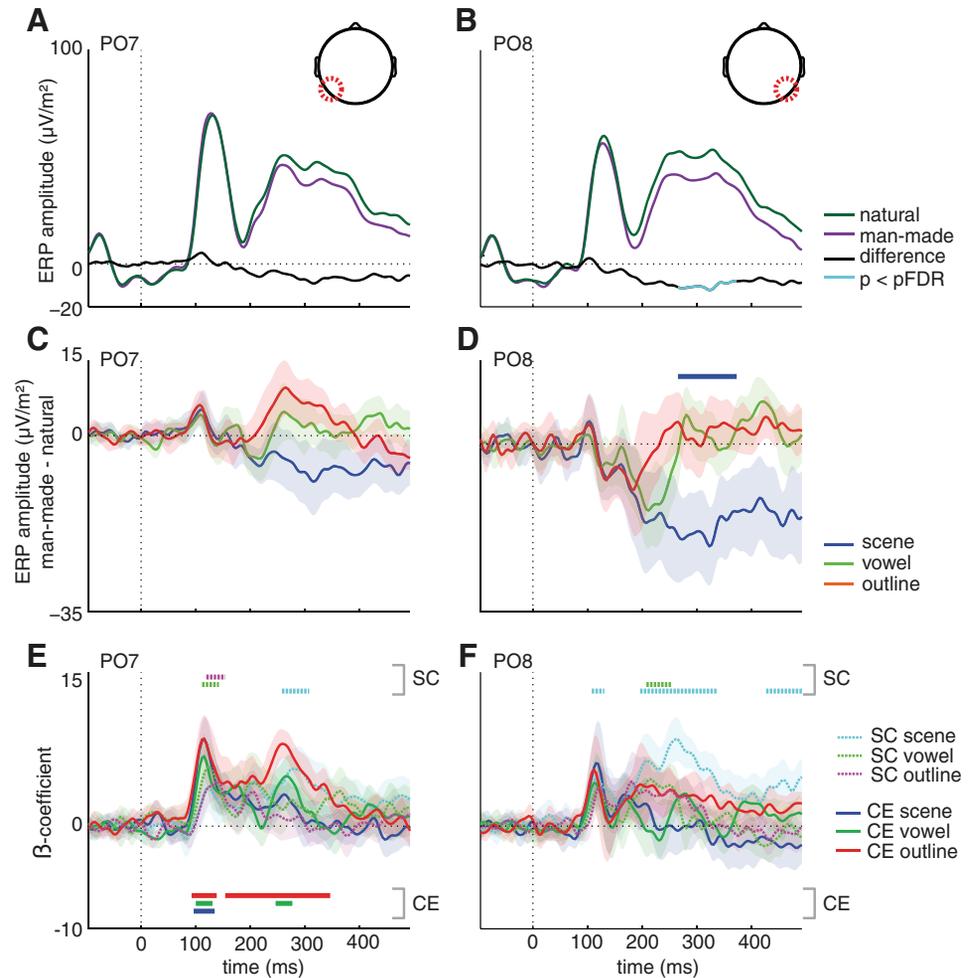


Fig. 9. Single-electrode ERP and regression results (Experiment 2) as in Fig. 8 but for parietal-occipital electrodes PO7 (A, C, and E) and PO8 (B, D, and F).

between experiments [$t(31) = 0.44$, $P = 0.67$]. The two experiments were conducted in the same room with the same screen settings. The only difference between the two experiments was the addition of the outline around the scene. We speculate that this additional stimulation may have negatively affected the speed of visual information processing. For example, the black and white peripheral outline may have selectively activated the magno-cellular pathway, diminishing its capacity to pick up coarse scene information (Delorme et al. 1999). The scene would then be relatively stronger, represented by the color-sensitive parvocellular pathway, which responds slightly slower (Maunsell and Newsome 1987). Alternatively, the presence of the high-contrast outline may have led to neural adaptation or contrast gain shifts (Huang and Dobkins 2005; Wagge and Olzak 2008), leading to slower processing of the scenes.

DISCUSSION

We examined how manipulations of attention affected scene information processing. Two independent experiments showed that, when a scene is presented in full view, but attention is directed to either a focal or a peripheral stimulus, global scene information is still present in early (<250 ms) evoked activity. This information was evident as a grand average difference in manmade vs. natural scenes that was indistinguishable between different attentional manipulations. Moreover, intact single-

trial sensitivity to global image statistics CE and SC indicated that the brain also processed information that is diagnostic of the degree of naturalness at the individual image level regardless of attentional locus.

At later time points, the enhanced attention gave rise to prolonged ERP signatures of scene processing that were more spread out across the scalp. Importantly, this enhanced processing had behavioral consequences; in both experiments, we observed better memory for scenes that had been categorized. Across the two experiments, larger negativity for manmade compared with natural scenes was only observed in the scene categorization task, suggesting that, although scene information is initially processed automatically, explicit categorization gives rise to deeper processing of this information.

In Experiment 2, we observed an unexpected inverted manmade/natural difference (more positive ERP amplitude for manmade scenes) for left hemisphere electrodes in the peripheral attention condition. However, subjects were as impaired at remembering scenes during this task as during the focal attention task. This suggests that enhanced later activity does not necessarily reflect enhanced processing of scene-specific information. Interestingly, unlike in Experiment 1, this inverse pattern of activity was associated with increased sensitivity to CE. Thus we speculate that, in this task, in which subjects had to discriminate the orientation of a high-contrast outline, neural sensitivity to contrast was enhanced, thereby giving rise to the

inverse difference wave, i.e., higher activity for manmade (which on average have higher CE values, see Fig. 1C) than natural scenes. SC, on the other hand, more closely followed the characteristic manmade/natural difference waves; sensitivity to SC was present in all tasks whenever ERP amplitude for natural scenes was higher than for manmade scenes, both early and late in time.

Thus, in line with the stronger correlation of SC than CE with behavioral naturalness rating (Groen et al. 2013), we here again find that SC is more tightly linked to the naturalness distinction because it selectively modulated late activity related to scene categorization. Together, these findings support a model of visual perception in which real-world scene information is extracted automatically in visual processing but flexibly processed to higher levels depending on task requirements.

Complex Perception Without Attention?

The attentional requirements of scene perception are intensely debated. Studies on dual-task and parallel image presentation have suggested that the brain constructs complex scene representations in the near absence of attention (Li et al. 2002; Rousselet et al. 2002; VanRullen et al. 2005). However, others have countered this claim, arguing that scene perception does require attention, given that humans exhibit considerable inattention blindness to natural scenes (Mack and Clarke 2012; Rensink 2001; Simons and Rensink 2005) and that scene perception is impaired under attentionally demanding rapid serial visual presentation or multiple object tracking tasks (Cohen et al. 2011; Evans and Treisman 2005).

To resolve this debate, it is necessary to understand the format of scene representation under different forms of attention (Rosenholtz et al. 2012a), i.e., what type of scene information can be processed to which levels of representation with and without attention. A recent fMRI study with naturalistic face stimuli showed that spatial attention manipulations affect visual responses in higher-level category-selective areas but not in early visual cortex (Kay et al. 2015), suggesting that low-level information processing is less susceptible to reduced attention. Similarly, EEG results have shown that, when subjects detect objects in natural scenes, very early sensory processing is not affected by the target vs. nontarget status of the scenes (VanRullen and Thorpe 2001); in that study, the task started to affect evoked responses from 150 ms onward. Another EEG experiment showed that segmentation of artificial textures can take place even when subjects do not attend to these textures (Scholte et al. 2006). Such early segmentation processes could reflect grouping of low-level features within the feed-forward sweep of visual activation (Lamme and Rolfsema 2000) in a “preattentive” parallel processing phase (Rousselet et al. 2004) in which a large amount of scene information is processed in a short period of time but not necessarily in much detail. In a similar vein, it has been proposed that the brain constructs a coarse “gist” representation via a separate processing pathway (Wolfe et al. 2011). Importantly, such a coarse sketch of a scene is thought to be sufficient for some semantic content extraction (Oliva 2005; Schyns and Oliva 1994), for example basic-level categorical labels like “beach” or “street scene” (Oliva and Torralba 2006) or global scene properties such as naturalness or spatial layout (Greene and Oliva 2009).

Indeed, despite the heated debate, most researchers in fact agree that coarse scene representations of “relatively low perceptual resolution” (Fize et al. 2005) may have “aspects that are systematically immune to attentional interference” (Cohen et al. 2011). The coarse nature of this representation may be the reason why it is more robust to attention manipulations than the artificial stimuli used in classic visual search (Rensink 2001; Treisman 2006), for which more detailed analysis is necessary (VanRullen et al. 2005). Potentially, this type of representation could even facilitate more selective binding processes (Oliva and Torralba 2007; Wolfe et al. 2011) and has thus been described as “nonselective,” “global,” or “statistical” (Alvarez and Oliva 2009; Oliva and Torralba 2007; Wolfe et al. 2011). However, the exact nature of this representation is unknown, and there has been very little investigation into any potential effects on neurophysiological signatures of early neural processing of global and/or statistical scene properties. Our findings provide initial steps toward more insight into this matter by providing the first neural evidence that processing of scene naturalness and local contrast statistics is only influenced by attention at later stages of visual processing.

A Role for Image Statistics in Scene Perception

Our results are consistent with the view that image statistics play a role in shaping neural representations (Balas et al. 2009; Freeman and Simoncelli 2011; Groen et al. 2012a, 2012b; Loschky et al. 2009; Scholte et al. 2009). We have argued previously that estimating CE and SC could be a useful step in visual processing because they index the fragmentation of the visual input, i.e., whether the scene is very cluttered and chaotic or strongly segmented and organized (Groen et al. 2013). This information is useful for estimating naturalness because natural scenes tend to be more fragmented. This, however, does not imply that these contrast statistics should be the only source of information relevant for a manmade/natural decision. Isolating the influence of these statistics would require comparing our results with categorization performance on manmade and natural scenes that do not differ in their underlying statistics. However, a complication of that experimental approach is that, because of the inherent correlation of scene statistics with category, such a subselection of scenes may no longer be an adequate sample of the real-world variance that the brain is tuned to and most likely optimized for (Fei-Fei et al. 2005; Felsen and Dan 2005; Olshausen and Field 1996; Vinje and Gallant 2000).

The early onsets of the ERP sensitivity to image statistics suggest that they mainly play a role in early visual processing, and it is likely that the format of scene representation changes substantially over the course of visual processing, transforming from an early, coarse representation closely tied to scene statistics toward a much richer and detailed representation at later stages of processing. Interestingly, however, we here still observed sensitivity to SC at relatively late time points (up to 300 ms). This suggests that scene statistics not only affect early visual representations, but can also be flexibly maintained depending on the task the subject is engaged in. Although SC is derived from local contrast information, it thus potentially plays a role in more high-level categorization and recognition processes. Indeed, low-level scene information is not necessarily completely irrelevant at later stages of visual processing.

For example, transcranial magnetic stimulation studies indicate that lower-level visual areas continue to play a role in scene processing beyond their initial responses (Koivisto et al. 2011), and a recent magnetoencephalography study showed that categorical object representations at later time points are still related to information encoded in V1 (Cichy et al. 2014).

On the other hand, there is evidence that, even early in visual processing, the extraction of scene information can be flexible. For example, a top-down task manipulation that required only image recognition rather than processing of specific scene information was capable of enhancing neural sensitivity to low-level cues (Delorme et al. 2004). It has also been shown that observers can selectively extract information present at specific spatial scales that are relevant for the current task (Malcolm et al. 2014; Oliva and Schyns 1997). The processing of information from specific spatial frequency bands has also been related to particular peaks in the event-related potential, with the P1 carrying mostly low spatial frequency information and the N1 carrying high spatial frequency information (Hansen et al. 2012). The scene categorization task in the current experiments did not require using high vs. low spatial frequencies per se. In fact, both of these are captured by the CE/SC model, which uses minimal reliable scale selection to extract contrast across multiple spatial scales (Ghebreab et al. 2009). Accordingly, the difference waves and the effects of image statistics we observed in the present study were present on multiple early sensory ERP peaks (P1, N1, extending into the P2 window). Previously, we found that spatial frequency-based image statistics did not predict ERP amplitude as well as CE/SC (Groen et al. 2012a, 2013), which was the main reason not to include an analysis of this type of information in the present study. Thus it remains unclear to what degree early neural activity may reflect specific spatial frequency information extraction dependent on the task requirements.

Effects of Spatial Location on the Attentional Time Course

It is clear that, in our paradigm, attention was reduced, but not completely removed from the scenes. We presented the scenes centrally in full view, and distracter stimuli were presented on or directly adjacent to the scenes. If the scenes were not visually processed at all, we would not have observed scene-evoked activity in the letter and outline conditions. The late attentional effects reported here could possibly shift to earlier time points in different experimental designs, for example, if the scene and distracter stimuli were more spatially separated, as in classic spatial cuing paradigms (Luck et al. 2000) or as in the study by Li et al. (2002).

However, present evidence argues against a large influence of spatial position in natural scene perception; ERPs are similar for scenes presented off center vs. centrally; even the scene location is unpredictable (Fize et al. 2005). Moreover, although scene categorization is generally better when scenes are presented at center, it remains above chance for scenes presented in the periphery (interestingly, with particular good performance for naturalness categorization; Boucart et al. 2013). Patients with impaired central vision can still categorize natural scenes (Tran et al. 2010), as can healthy subjects presented with peripheral information only (Larson and Loschky 2009). Thus scene categorization is possible based on peripheral vision, but it is unclear how this ability is affected by atten-

tional manipulations. Future experiments can establish what type of scene information can be detected in ERPs when attention is directed centrally while scenes are presented in the periphery. Possibly, image statistics play an even stronger role in peripheral vision because perceptually indistinguishable scenes can be constructed by matching peripheral statistics (Freeman and Simoncelli 2011; Rosenholtz et al. 2012b).

Hemispheric Asymmetry in Late Effects

The fact that, in Experiment 2, a task-dependent inverted ERP effect and sensitivity to contrast energy manifested in the left but not the right hemispheres is intriguing. A popular theory on hemispheric asymmetries in visual processing posits that the left and right hemispheres are attentionally biased toward local and global processing, respectively (Heinze et al. 1998; Robertson and Lamb 1991; Yamaguchi et al. 2000). With the use of compound stimuli (typically Navon's hierarchical letters, Navon 1977), it has been found that this local/global asymmetry can be modulated by many factors, including perceptual salience, the degree of response conflict, as well as spatial frequency and the level of contrast (Fink et al. 1998; Han et al. 2003; Volberg and Hübner 2004). Interestingly, it appears that greater contrast differences between local and global stimulus properties elicit larger hemispheric asymmetry by biasing the left hemisphere more toward local processing (Jiang and Han 2005).

Thus one possibility is that the high-contrast outline specifically elicited more local processing of the local boundaries of the outlines in the left hemisphere. However, an alternative explanation might be that the more difficult outline task elicited more response conflict between the scene and the outline leading to a stronger "division of labor" between hemispheres (Malinowski et al. 2002). In general, caution is warranted in these post hoc interpretations, as well as for the reason that most previous literature pertains to artificial stimuli. There is some evidence that these hemispheric biases may also exist for more naturalistic stimuli such as objects (Hübner and Studer 2009) or scenes that have been filtered to contain only low or high spatial frequencies (Peyrin et al. 2005, 2010), but it is unclear how the hemispheric differences in attentional bias toward local and global information may affect the neural representation of intact, real-world scenes and their scene statistics.

Future Directions: Exploring a Potential Role for Perceptual Load

In Experiment 1, despite the difference in task difficulty between the vowel and 2-back letter task, we found an equal decrease in scene memory. Moreover, the attentional effects on the scene-evoked activity were highly similar between these two tasks, with a clear lack of interference by the letter tasks before 250 ms and a clear lack of the scene representation beyond that time point. This is consistent with the idea that processing of irrelevant visual input is affected by increasing perceptual but not working memory load (Handy et al. 2001; Lavie 1995; Yi et al. 2004). In our study, even though the central letters were small and rapidly presented, they were clearly visible, suggesting that their perceptual load was not extremely high. It would be interesting to examine to what extent the ERP differences between manmade and natural

scenes are still present when the letter task is made more perceptually difficult, e.g., by blurring the letters. In Experiment 2, the outline task was more difficult than the other tasks, suggesting that it was more perceptually demanding. However, the early ERP differences were still intact. It has been suggested that, in natural scene perception, early competition for attention occurs only if target and irrelevant information are spatially overlapping (Lavie 1995; VanRullen et al. 2005). Here, the outline was spatially adjacent to the scene but was visually very distinct; it remains unclear to what extent competition may have played a role here.

Conclusion

We provide electrophysiological evidence that the extraction of information from real-world scenes remains intact when attention to the scene is reduced. However, increased attention does enhance real-world scene information processing at later processing stages.

ACKNOWLEDGMENTS

We thank Rianne Penterman for collecting the data for Experiment 1.

GRANTS

This work is part of the Research Priority Program “Brain & Cognition” at the University of Amsterdam and was supported by an Advanced Investigator Grant from the European Research Council to V. A. F. Lamme and the Dutch national public-private research program COMMIT to S. Ghebreab.

DISCLOSURES

No conflicts of interest, financial or otherwise, are declared by the authors.

AUTHOR CONTRIBUTIONS

Author contributions: I.I.A.G., S.G., V.A.F.L., and H.S.S. conception and design of research; I.I.A.G. performed experiments; I.I.A.G. analyzed data; I.I.A.G., S.G., and H.S.S. interpreted results of experiments; I.I.A.G. prepared figures; I.I.A.G. drafted manuscript; I.I.A.G., S.G., V.A.F.L., and H.S.S. edited and revised manuscript; I.I.A.G., S.G., V.A.F.L., and H.S.S. approved final version of manuscript.

REFERENCES

- Alvarez GA, Oliva A. Spatial ensemble statistics are efficient codes that can be represented with reduced attention. *Proc Natl Acad Sci USA* 106: 7345–7350, 2009.
- Balas B, Nakano L, Rosenholtz R. A summary-statistic representation in peripheral vision explains visual crowding. *J Vis* 9: 13.1–13.18, 2009.
- Bieniek MM, Frei LS, Rousselet GA. Early ERPs to faces: aging, luminance, and individual differences. *Front Psychol* 4: 268, 2013.
- Bonin V, Mante V, Carandini M. The suppressive field of neurons in lateral geniculate nucleus. *J Neurosci* 25: 10844–10856, 2005.
- Boucart M, Moroni C, Thibaut M, Szaffarczyk S, Greene M. Scene categorization at large visual eccentricities. *Vision Res* 86: 35–42, 2013.
- Brady N, Field DJ. Local contrast in natural images: normalisation and coding efficiency. *Perception* 29: 1041–1055, 2000.
- Braun J. Natural scenes upset the visual appercept. *Trends Cogn Sci* 7: 7–9, 2003.
- Cichy RM, Pantazis D, Oliva A. Resolving human object recognition in space and time. *Nat Neurosci*: 1–10, 2014.
- Cohen MA, Alvarez GA, Nakayama K. Natural-scene perception requires attention. *Psychol Sci* 22: 1165–1172, 2011.
- Delorme A, Richard G, Fabre-Thorpe M. Rapid processing of complex natural scenes: A role for the magnocellular visual pathways? *Neurocomputing* 26: 663–670, 1999.
- Delorme A, Rousselet GA, Macé MJM, Fabre-Thorpe MM. Interaction of top-down and bottom-up processing in the fast visual analysis of natural scenes. *Cogn Brain Res* 19: 103–113, 2004.
- Elder JH, Zucker SW. Local scale control for edge detection and blur estimation. *IEEE Trans Pattern Anal Mach Intel* 20: 699–716, 1998.
- Evans KK, Treisman A. Perception of objects in natural scenes: is it really attention free? *J Exp Psychol Hum Percept Perform* 31: 1476–1492, 2005.
- Fei-Fei L, VanRullen R, Koch C, Perona P. Why does natural scene categorization require little attention? Exploring attentional requirements for natural and synthetic stimuli. *Vis Cogn* 12: 893–924, 2005.
- Felsen G, Dan Y. A natural approach to studying vision. *Nat Neurosci* 8: 1643–1646, 2005.
- Fink GR, Marshall JC, Halligan PW, Dolan RJ. Hemispheric asymmetries in global/local processing are modulated by perceptual salience. *Neuropsychologia* 37: 31–40, 1998.
- Fize D, Fabre-Thorpe M, Richard G, Doyon B, Thorpe SJ. Rapid categorization of foveal and extrafoveal natural images: associated ERPs and effects of lateralization. *Brain Cogn* 59: 145–158, 2005.
- Freeman J, Simoncelli EP. Metamers of the ventral stream. *Nat Neurosci* 14: 1195–1201, 2011.
- Ghebreab S, Smeulders AWM, Scholte HS, Lamme VAF. A biologically plausible model for rapid natural image identification. In: *Advances in Neural Information Processing Systems*. Vancouver, Canada: NIPS, 2009, p. 629–637.
- Gratton C. A new method for off-line removal of ocular artifact. *Electroencephalogr Clin Neurophysiol* 55: 468–484, 1983.
- Greene MR, Oliva A. The briefest of glances: the time course of natural scene understanding. *Psychol Sci* 20: 464–472, 2009.
- Groen IIA, Ghebreab S, Lamme VA, Scholte HS. Spatially pooled contrast responses predict neural and perceptual similarity of naturalistic image categories. *PLoS Comput Biol* 8: e1002726, 2012a.
- Groen IIA, Ghebreab S, Lamme VAF, Scholte HS. Low-level contrast statistics are diagnostic of invariance of natural textures. *Front Comput Neurosci* 6: 34, 2012b.
- Groen IIA, Ghebreab S, Prins H, Lamme VAF, Scholte HS. From image statistics to scene gist: evoked neural activity reveals transition from low-level natural image structure to scene category. *J Neurosci* 33: 18814–18824, 2013.
- Han S, Yund EW, Woods DL. An ERP study of the global precedence effect: The role of spatial frequency. *Clin Neurophysiol* 114: 1850–1865, 2003.
- Handy TC, Soltani M, Mangun GR. Perceptual load and visuospatial processing: event-related potentials reveal sensory-level selection. *Psychol Sci* 12: 213–218, 2001.
- Hansen BC, Johnson AP, Ellemberg D. Different spatial frequency bands selectively signal for natural image statistics in the early visual system. *J Neurophysiol* 108: 2160–2172, 2012.
- Heinze HJ, Hinrichs H, Scholz M, Burchert W, Mangun GR. Neural mechanisms of global and local processing. A combined PET and ERP study. *J Cogn Neurosci* 10: 485–498, 1998.
- Huang L, Dobkins KR. Attentional effects on contrast discrimination in humans: evidence for both contrast gain and response gain. *Vision Res* 45: 1201–1212, 2005.
- Hübner R, Studer T. Functional hemispheric differences for the categorization of global and local information in naturalistic stimuli. *Brain Cogn* 69: 11–18, 2009.
- Jiang Y, Han S. Neural mechanisms of global/local processing of bilateral visual inputs: An ERP study. *Clin Neurophysiol* 116: 1444–1454, 2005.
- Johnson JS, Olshausen BA. Timecourse of neural signatures of object recognition. *J Vis* 3: 499–512, 2003.
- Kastner S, Pinsk MA. Visual attention as a multilevel selection process. *Cogn Affect Behav Neurosci* 4: 483–500, 2004.
- Kay KN, Weiner KS, Grill-Spector K. Attention reduces spatial uncertainty in human ventral temporal cortex. *Curr Biol* 25: 595–600, 2015.
- Koenderink JJ, van de Grind WA, Bouman MA. Opponent color coding: A mechanistic model and a new metric for color space. *Kybernetik* 10: 78–98, 1972.
- Koivisto M, Railo H, Revonsuo A, Vanni S, Salminen-Vaparanta N. Recurrent processing in V1/V2 contributes to categorization of natural scenes. *J Neurosci* 31: 2488–2492, 2011.
- Kravitz DJ, Saleem KS, Baker CI, Ungerleider LG, Mishkin M. The ventral visual pathway: an expanded neural framework for the processing of object quality. *Trends Cogn Sci* 17: 26–49, 2013.
- Lamme V, Roelfsema P. The distinct modes of vision offered by feedforward and recurrent processing. *Trends Neurosci* 23: 571–579, 2000.

- Larson AM, Loschky LC.** The contributions of central versus peripheral vision to scene gist recognition. *J Vis* 9: 1–16, 2009.
- Lavie N.** Perceptual load as a necessary condition for selective attention. *J Exp Psychol Hum Percept Perform* 21: 451–468, 1995.
- Li FF, VanRullen R, Koch C, Perona P.** Rapid natural scene categorization in the near absence of attention. *Proc Natl Acad Sci USA* 99: 9596–9601, 2002.
- Loschky LC, Hansen BC, Sethi A, Pydimarri TN.** The role of higher order image statistics in masking scene gist recognition. *Atten Percept Psychophys* 71: 481–489, 2009.
- Luck SJ.** *An Introduction to the Event-Related Potential Technique.* Cambridge, MA: MIT, 2005.
- Luck S, Woodman G, Vogel E.** Event-related potential studies of attention. *Trends Cogn Sci* 4: 432–440, 2000.
- Mack A, Clarke J.** Gist perception requires attention. *Vis Cogn* 20: 300–327, 2012.
- Macmillan NA, Creelman CD.** *Detection Theory: A User's Guide.* New York, NY: Cambridge University, 1991.
- Malcolm GL, Nuthmann A, Schyns PG.** Beyond gist: Strategic and incremental information accumulation for scene categorization. *Psychol Sci* 25: 1087–1097, 2014.
- Malinowski P, Hübner R, Keil A, Gruber T.** The influence of response competition on cerebral asymmetries for processing hierarchical stimuli revealed by ERP recordings. *Exp Brain Res* 144: 136–139, 2002.
- Marois R, Ivanoff J.** Capacity limits of information processing in the brain. *Trends Cogn Sci* 9: 296–305, 2005.
- Maunsell JHR, Newsome WT.** Visual processing in monkey extrastriate cortex. *Annu Rev Neurosci* 10: 363–401, 1987.
- Navon D.** Forest before trees: The precedence of global features in visual perception. *Cogn Psychol* 9: 353–383, 1977.
- Oliva A.** Gist of the scene. In: *Neurobiology of Attention.* New York, NY: Elsevier, 2005, p. 251–257.
- Oliva A, Schyns PG.** Coarse blobs or fine edges? Evidence that information diagnosticity changes the perception of complex visual stimuli. *Cogn Psychol* 34: 72–107, 1997.
- Oliva A, Torralba A.** Building the gist of a scene: the role of global image features in recognition. *Prog Brain Res* 155: 23–36, 2006.
- Oliva A, Torralba A.** The role of context in object recognition. *Trends Cogn Sci* 11: 520–527, 2007.
- Olshausen BA, Field DJ.** Emergence of simple-cell receptive field properties by learning a sparse code for natural images. *Nature* 381: 607–610, 1996.
- Peelen MV, Kastner S.** Attention in the real world: toward understanding its neural basis. *Trends Cogn Sci* 18: 242–250, 2014.
- Perrin F.** Spherical splines for scalp potential and current density mapping. *Electroencephalogr Clin Neurophysiol* 72: 184–187, 1989.
- Peyrin C, Schwartz S, Seghier M, Michel C, Landis T, Vuilleumier P.** Hemispheric specialization of human inferior temporal cortex during coarse-to-fine and fine-to-coarse analysis of natural visual scenes. *Neuroimage* 28: 464–473, 2005.
- Peyrin C, Michel CM, Schwartz S, Thut G, Seghier M, Landis T, Marandaz C, Vuilleumier P.** The neural substrates and timing of top-down processes during coarse-to-fine categorization of visual scenes: a combined fMRI and ERP study. *J Cogn Neurosci* 22: 2768–2780, 2010.
- Potter MC.** Meaning in visual search. *Science* 187: 965–966, 1975.
- Rensink RA.** Change blindness: Implications for the nature of visual attention. In: *Vision & Attention*, edited by Jenkin M, Harris L. New York, NY: Springer, 2001, p. 169–188.
- Rieger JW, Gegenfurtner KR, Koechy N, Heinze HJ, Grueschow M.** BOLD responses in human V1 to local structure in natural scenes: Implications for theories of visual coding. *J Vis* 13: 1–15, 2013.
- Robertson LC, Lamb MR.** Neuropsychological contributions to theories of part/whole organization. *Cogn Psychol* 23: 299–330, 1991.
- Rosenholtz R, Huang J, Ehinger KA.** Rethinking the role of top-down attention in vision: Effects attributable to a lossy representation in peripheral vision. *Front Psychol* 3: 1–15, 2012a.
- Rosenholtz R, Huang J, Raj A, Balas BJ, Ilie L.** A summary statistic representation in peripheral vision explains visual search. *J Vis* 12: 14, 2012b.
- Rousselet GA, Fabre-Thorpe M, Thorpe SJ.** Parallel processing in high-level categorization of natural images. *Nat Neurosci* 5: 629–630, 2002.
- Rousselet GA, Thorpe SJ, Fabre-Thorpe M.** How parallel is visual processing in the ventral pathway? *Trends Cogn Sci* 8: 363–370, 2004.
- Scholte HS, Witteveen SC, Spekreijse H, Lamme VAF.** The influence of inattention on the neural correlates of scene segmentation. *Brain Res* 1076: 106–115, 2006.
- Scholte HS, Ghebreab S, Waldorp L, Smeulders AWM, Lamme VAF.** Brain responses strongly correlate with Weibull image statistics when processing natural images. *J Vis* 9: 1–15, 2009.
- Schyns PG, Oliva A.** From blobs to boundary edges: Evidence for time- and spatial-scale-dependent scene recognition. *Psychol Sci* 5: 195–200, 1994.
- Simons DJ, Rensink R.** Change blindness: past, present, and future. *Trends Cogn Sci* 9: 16–20, 2005.
- Torralba A, Oliva A.** Statistics of natural image categories. *Netw Comput Neural Syst* 14: 391–412, 2003.
- Tran THC, Rambaud C, Despretz P, Boucart M.** Scene perception in age-related macular degeneration. *Invest Ophthalmol Vis Sci* 51: 6868–6874, 2010.
- Treisman A.** How the deployment of attention determines what we see. *Vis Cogn* 14: 411–443, 2006.
- Ungerleider LG, Mishkin M.** Two cortical visual systems. In: *Analysis of Visual Behavior*, edited by Ingle DI, Goodale MA, Mansfield RJ. Cambridge, MA: MIT Press, 1982, p. 549–586.
- VanRullen R, Thorpe S.** The time course of visual processing: from early perception to decision-making. *J Cogn Neurosci* 13: 454–461, 2001.
- VanRullen R, Reddy L, Fei-Fei L.** Binding is a local problem for natural objects and scenes. *Vision Res* 45: 3133–3144, 2005.
- Vinje WE, Gallant JL.** Sparse coding and decorrelation in primary visual cortex during natural vision. *Science* 287: 1273–1276, 2000.
- Volberg G, Hübner R.** On the role of response conflicts and stimulus position for hemispheric differences in global/local processing: An ERP study. *Neuropsychologia* 42: 1805–1813, 2004.
- Wagge JR, Olzak LA.** Contributions of contrast gain and response gain in pattern masking. *J Vis* 8: 89, 2008.
- Wolfe JM.** Visual search. In: *Attention*, edited by Pashler H. London, UK: London College University, 1998, p. 13–73.
- Wolfe JM, Horowitz TS, Michod KO.** Is visual attention required for robust picture memory? *Vision Res* 47: 955–964, 2007.
- Wolfe JM, Võ MLH, Evans KK, Greene MR.** Visual search in scenes involves selective and nonselective pathways. *Trends Cogn Sci* 15: 77–84, 2011.
- Yamaguchi S, Yamagata S, Kobayashi S.** Cerebral asymmetry of the “top-down” allocation of attention to global and local features. *J Neurosci* 20: RC72, 2000.
- Yi DJ, Woodman GF, Widders D, Marois R, Chun MM.** Neural fate of ignored stimuli: dissociable effects of perceptual and working memory load. *Nat Neurosci* 7: 992–996, 2004.
- Zhu SC, Mumford D.** Prior learning and Gibbs reaction-diffusion. *IEEE Trans Pattern Analysis Machine Intel* 19: 1236–1250, 1997.