Published in final edited form as:

Brain Res. 2011 June 29; 1398: 40–54. doi:10.1016/j.brainres.2011.04.029.

Contextual Influences on Rapid Object Categorization in Natural Scenes

Hsin-Mei Sun¹, Stephanie L. Simon-Dack², Robert D. Gordon¹, and Wolfgang A. Teder¹

¹ Department of Psychology, North Dakota State University, Fargo, ND 58105, USA

² Department of Psychological Science, Ball State University, Muncie, IN 47306, USA

Abstract

The current study aimed to investigate the effects of scene context on rapid object recognition using both behavioral and electrophysiological measures. Participants performed an animal/nonanimal go/no-go categorization task in which they had to decide whether or not a flashed scene contained an animal. Moreover, the influence of scene context was manipulated either by retaining, deleting, or phase-randomizing the original scene background. The results of Experiments 1 and 2 showed that participants responded more accurately and quickly to objects appearing with their original scene backgrounds. Moreover, the event-related potential (ERP) data obtained from Experiment 2 showed that the onset latency of the frontal go/no-go ERP difference was delayed for objects appearing with phase-randomized scene backgrounds compared to objects appearing with their original scene backgrounds, providing direct evidence that scene context facilitates object recognition. Additionally, an increased frontal negativity along with a decreased late positive potential for processing objects presented in meaningless scene backgrounds suggest that the categorization task becomes more demanding when scene context is eliminated. Together, the results of the current study are consistent with previous research showing that scene context modulates object processing.

Keywords

object categorization; natural scenes; context effects; event-related potentials (ERPs)

1. Introduction

Target detection in natural scenes can be performed successfully even when the stimulus presentation time is shorter than a single glance (e.g., within one fixation). For example, Potter (1975) gave participants a brief description of the main objects or event in a scene (e.g., a boat, two men drinking beer) and then asked them to detect the target picture in a sequence of rapidly presented scenes. The results showed that participants could detect more than 70 percent of the targets when the sequences were presented at the rapid rate of 125 ms per picture, demonstrating that less than 125 ms is needed for recognizing the content of a complex image (see also Potter, 1976). Similarly, Intraub (1981) asked participants to detect a verbally specified target (e.g., a rose) while viewing a rapid sequence of pictures, and the

Address correspondence to: Hsin-Mei Sun, Department of Psychology, North Dakota State University, Fargo, ND 58105, USA, Phone: (701) 231-8622, Fax: (701) 231-8426, mei.sun@ndsu.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

results showed that more than 70 percent of the targets cued by specific name could be detected at the presentation rate of 114 ms per picture. These findings reveal that the detection of target objects in natural scenes can be achieved efficiently.

Given that objects can be categorized efficiently even when they are embedded in rapidly presented scenes, an interesting question is whether scene context contributes to such remarkable performance. Earlier studies using line-drawn pictures have shown that when participants are presented with a scene depicting a certain context, objects that are consistent with that context are recognized more easily than objects that would not be expected in that context. For example, the context of a kitchen can facilitate recognition of a loaf of bread in comparison to a drum (Palmer, 1975). In addition, observers are more likely to attend to semantically inconsistent objects (e.g., a fire hydrant in a bedroom) during free viewing, probably because these objects are relatively difficult to identify in an inappropriate context (Gordon, 2004, 2006). Finally, objects are recognized more efficiently when they appear in a semantically consistent background (Biederman, Mezzanotte, & Rabinowitz, 1982; Boyce & Pollatsek, 1992; Boyce, Pollatsek, & Rayner, 1989).

More recently, research using naturalistic color photographs has further shown that the effect of scene context on object processing could be measured by recording event-related potentials (ERPs). For example, Ganis and Kutas (2003) presented participants with a fixation cross, followed by a scene (e.g., soccer players in a soccer field). The location of the fixation cross varied from trial to trial and served as a pre-cue to indicate the location of an upcoming target object. After 300 ms, a semantically congruent (e.g., a soccer) or incongruent (e.g., a toilet paper roll) object appeared at the cued location and was shown together with the scene for 300 ms; participants were asked to identify the target object that appeared at the cued location. Ganis and Kutas (2003) showed that the processing of objects embedded in an incongruent context is associated with a larger N390, which is a negative-going ERP component occurs between 300 and 500 ms after stimulus presentation. Given that the N390 scene congruity effect is similar to the N400 sentence congruity effect that is typically found for a verbal stimulus that violates the semantic context created by preceding stimuli (e.g., Kutas & Hillyard, 1980), Ganis and Kutas suggested that the N390 scene congruity effect reflects the influence of scene context on object processing at the level of semantic analysis. The N390 scene congruity effect was replicated in a recent study using the pre-cue procedure but presenting a semantically congruent or incongruent object with a scene simultaneously for 1000 ms (Mudrik, Lamy, & Deouell, 2010). Similar to studies of the scene context effect on object recognition, research investigating how emotional scenes affect the recognition of facial expressions has shown that the N170 response to faces is larger for fearful faces in a fearful context, which provides further evidence for the scene-object congruency effect (e.g., de Gelder et al., 2006; Righart & de Gelder, 2006).

Recent studies have also shown that scene background is able to affect object processing even when an image is glimpsed briefly. Davenport and Potter (2004), for example, had participants report the name of an object embedded in a rapidly presented (80 ms) scene and showed that participants reported objects more accurately when they appeared with a consistent background than when they appeared with an inconsistent background. Joubert, Fize, Rousselet and Fabre-Thorpe (2008, Experiment 2) reported similar results by using an animal/non-animal go/no-go categorization task in which participants had to decide whether a briefly presented (26 ms) scene contained an animal. Similar to Davenport and Potter's (2004) manipulations, objects were pasted into various scene backgrounds to create congruent or incongruent object-scene combinations. The results showed that participants' performance was less accurate and slower when the target object was embedded in a semantically inconsistent scene background, such as an elephant appearing in a city scene.

Therefore, these findings support the hypothesis that scene context affects object processing even when an image is presented briefly.

However, there are some potential concerns with the stimulus manipulations for studies examining contextual influences on object recognition by pasting objects into new scene backgrounds. Joubert et al. (2008, Experiment 1), for example, observed that participants' categorization performance was impaired when foreground objects (e.g., a bicycle, a tiger) were cut from their original scene background and then pasted into new congruent backgrounds. That is, participants showed less accuracy and slower reaction time when they viewed a tiger that was cut from its original forest scene background and pasted into a mountain stream scene background, even though the new background was also consistent with the object's identity. This "pasting effect" might be due to changes in the local physical features (illumination and shadows) at the object-scene boundary when an object is pasted into a new background (Joubert et al., 2008).

To control the potential interference caused by pasting objects into new scene backgrounds, Davenport and Potter (2004) and Joubert et al. (2008, Experiment 2) had all stimuli contain a pasted object. That is, an object was segmented from its original scene background and then pasted into different scene backgrounds to create semantically congruent and incongruent pictures. In doing so, however, the potential problems with such stimulus manipulations still exist (e.g., incoherent illumination and shadows between the pasted object and its new scene background). Moreover, the segmented object may have a different spatial resolution than its new background, so that a high spatial resolution object image might be perceived as more salient if it is placed in a low spatial resolution scene background. Additionally, certain types of relations that characterize a scene, such as relative scales and supports (Biederman et al., 1982), may be violated easily when introducing a segmented object to a new background. For example, the perceived size of an object might change according to the perspective of the current background. If the perspectives of the two backgrounds are quite different, a cup copied from a kitchen scene to a living room scene may result in the cup looking unnaturally small or large.

The first goal of the current study, therefore, was to examine the influence of scene context on rapid object categorization while avoiding the pasting effect. In Experiment 1a, Participants were asked to perform an animal/nonanimal go/no-go categorization task in which they had to respond to animals appearing in briefly presented images. In addition, the presence of an object's original background information, rather than the congruency between an object and its background information, was manipulated to avoid the aforementioned pasting effect. One potential concern with this manipulation, however, is that recognition of an isolated object might benefit from its clear contour when it is presented alone on a blank background (e.g., Davenport & Potter, 2004). Therefore, in the present study, an isolated object was not segmented from the original image in the background absence condition. Instead, the object was cropped and embedded in a background in which the remainder of the image was either deleted or phase randomized. In doing so, the issue of object-background segmentation process was controlled and the availability of scene context was reduced to a minimum in the blank or phase-randomized background condition. If scene context affects rapid object recognition, participants should have better performance in categorizing objects appearing with the original scene backgrounds.

Experiment 1b was motivated by the fact that when an object was presented in a blank or phase-randomized background in Experiment 1a, it was surrounded by a high-contrast and high frequency box, which might have the potential to produce lateral masking. To evaluate whether the sharp edges of the box would impair participants' categorization performance

and thus confound the context effects, we applied Gaussian blur to smooth the edges of the box surrounding an object in the blank and phase-randomized background conditions. If Experiments 1a and 1b yield the same pattern of results, we could rule out the potential concern of a lateral masking effect caused by the clear borderlines of the box encompassing the object.

More importantly, the current study also aimed to examine how scene context affects the neural processes related to rapid object recognition in natural scenes. Research using electrophysiological measurement has shown that some form of high-level visual representation can be accessed rapidly and thus enable participants to categorize objects in briefly presented images (e.g., Rousselet, Fabre-Thorpe, & Thorpe, 2002; Thorpe, Fize, & Marlot, 1996; VanRullen & Thorpe, 2001). For example, Thorpe et al. (1996) asked participants to perform an animal/nonanimal go/no-go categorization task in which participants had to respond to a briefly presented (20 ms) natural scene if it contained an animal. Despite the complexity and the very short presentation times of the images, the results showed that participants were able to detect the presence of an animal with high accuracy and fast reaction time. Additionally, ERPs elicited by target (image with animals) and distractor (image without animals) pictures started to diverge at 150 ms after stimulus onset in the frontal region, suggesting that differential processing of target and distractor pictures takes no longer than 150 ms within the brain. Therefore, an interesting question arises as to whether scene context is able to modulate the time course of rapid object recognition as early as the frontal go/no-go ERP difference. For example, if scene context influences the early stages of object processing, the onset latency of the frontal go/no-go ERP difference should be affected by the availability of scene background information.

Experiment 2 was conducted to test the possibility mentioned above. Participants performed an animal/nonanimal go/no-go categorization task while their ERPs were recorded at the same time. The effect of scene context on object categorization was manipulated by either maintaining or phase-randomizing an object's original scene background. Note that we used only the phase-randomized scene background condition to test the effect of deleting scene background information on rapid object categorization in Experiment 2. One advantage of using phase-randomized scene backgrounds is that the procedure of phase randomization only changes an image's phase structure, but preserves other stimulus characteristics, such as overall luminance and spatial frequency. Therefore, phase-randomized backgrounds serve as better experimental stimuli than blank gray backgrounds in the current experiment, because observed differences in early ERP components could not be attributed to low-level differences in the experimental stimuli. As mentioned earlier, if scene context influences the early stages of object processing, one would expect to observe a modulating effect of scene background on the onset latency of the frontal go/no-go ERP difference. In particular, if scene context facilitates the early stages of object processing, the onset latency of the frontal go/no-go ERP difference should be shorter for objects appearing in their original backgrounds compared to objects appearing in phase-randomized backgrounds.

In addition to the onset latency of the frontal go/no-go ERP difference, two ERP components, the frontal negativity and the late positive potential, were also assessed based on previous findings regarding visual object recognition (e.g., Bokura, Yamaguchi, & Kobayashi, 2001; Codispoti, Ferrari, Junghöfer, & Schupp, 2006; Eimer, 1993; Falkenstein, Hoormann, & Hohnsbein, 1999; Ferrari, Codispoti, Cardinale, & Bradley, 2008; Kok, 1997, 2001). The frontal negativity occurs approximately 200 ms after stimulus onset, and is typically larger on no-go than on go trials. Research has suggested that the enhanced frontal negativity observed in a go/no-go task reflects processes involved in the inhibition of motor responses (Bokura et al., 2001; Eimer, 1993; Falkenstein et al., 1999). Therefore, it was expected that a larger frontal negativity would be observed for objects appearing without the

original scene backgrounds, because the lack of scene context would make object recognition a more demanding process and thus more inhibitory effort would be needed to suppress the execution of responses before making a correct categorization decision.

In addition, previous studies have shown that a late positive potential, which occurs approximately 300 ms after stimulus onset over the centro-parietal recording sites, is larger for target stimuli than for nontarget stimuli, suggesting that more attentional resources are devoted to targets during object categorization processes (Codispoti et al., 2006; Ferrari et al., 2008). Moreover, the amplitude of the late positive potential is correlated with the efficiency of information updating and object processing (for a review, see Kok, 1997, 2001). Therefore, it was expected that an enhancement in the late positive potential amplitude would be observed for objects appearing with their original scene backgrounds, reflecting more efficient processing for target objects embedded in scenes.

2. Results

2.1 Experiment 1

In Experiment 1a, we had participants perform an animal/nonanimal go/no-go categorization task in which they had to make a response each time a flashed (20 ms) image contained an animal and withhold their response otherwise. In addition, the influence of scene context on object processing was tested by maintaining, deleting, or phase-randomizing an object's original scene background. Note that in the latter two conditions, an object was cropped and embedded in a blank or phase-randomized background so the object was surrounded by a box with sharp edges. To examine whether the sharp edges of the box might cause lateral masking and thus confound the context effect, we further conducted Experiment 1b, which was identical to Experiment 1a, except that we blurred the edges of the box that surrounded an object in the blank and phase-randomized background conditions.

Tables 1A and 1B show the mean accuracy and median reaction times in each of the experimental conditions from Experiments 1a and 1b, respectively. Note that the accuracy measures for animals are correct go responses (hits), whereas the accuracy measures for vehicles are correct no-go responses (correct rejections). The pattern of results was the same across the two experiments. Participants were very efficient at performing the animal/nonanimal go/no-go categorization task. The mean accuracy was above 80% for all six experimental conditions in both experiments. Two-way (object category $\times$ scene background) ANOVAs performed on the mean accuracy showed a significant main effect of object category: Experiment 1a, $F(1, 17) = 33.432$, $MSE = .015$, $p = .000$, and Experiment 1b, $F(1, 19) = 37.698$, $MSE = .006$, $p = .000$. That is, participants' overall accuracy was higher for the animal category. The main effect of scene background was also significant: Experiment 1a, $F(2, 34) = 4.158$, $MSE = .001$, $p = .024$, and Experiment 1b, $F(2, 38) = 4.197$, $MSE = .002$, $p = .023$. Planned comparisons showed that objects presented with their original scene backgrounds were reported more accurately than objects presented with either blank scene backgrounds: Experiment 1a, $F(1, 34) = 7.200$, $MSE = .001$, $p = .011$, and Experiment 1b, $F(1, 38) = 4.449$, $MSE = .002$, $p = .042$; or phase-randomized scene backgrounds: Experiment 1a, $F(1, 34) = 6.498$, $MSE = .001$, $p = .016$, and Experiment 1b, $F(1, 38) = 7.056$, $MSE = .002$, $p = .012$. However, there were no differences in accuracy between objects presented with blank scene backgrounds and objects presented with phase-randomized scene backgrounds: Experiment 1a, $F(1, 34) = .018$, $MSE = .001$, $p = .894$, and Experiment 1b, $F(1, 38) = .299$, $MSE = .002$, $p = .588$. Finally, there was no interaction between object category and scene background: Experiment 1a, $F(2, 34) = 1.617$, $MSE = .001$, $p = .213$, and Experiment 1b, $F(2, 38) = .729$, $MSE = .002$, $p = .489$.

Participants were also able to respond to the stimuli rapidly on correct go trials. One-way ANOVAs performed on the median reaction times for the animal category revealed a significant effect of scene background on target detection: Experiment 1a, $F(2, 34) = 8.912$, $MSE = 43.758$, $p = .001$, and Experiment 1b, $F(2, 38) = 11.973$, $MSE = 44.988$, $p = .000$. Planned comparisons revealed that the median reaction times for responding to animals appearing with their original scene backgrounds was significantly faster than to animals appearing with either blank scene backgrounds: Experiment 1a, $F(1, 34) = 22.763$, $MSE = 43.758$, $p = .000$, and Experiment 1b, $F(1, 38) = 23.594$, $MSE = 44.988$, $p = .000$; or phase-randomized scene backgrounds: Experiment 1a, $F(1, 34) = 30.188$, $MSE = 43.758$, $p = .000$, and Experiment 1b, $F(1, 38) = 44.857$, $MSE = 44.988$, $p = .000$. However, the median reaction times for responding to animals in the blank scene background condition did not differ significantly from the median reaction times for responding to animals in the phase-randomized scene background condition: Experiment 1a, $F(1, 34) = 0.523$, $MSE = 43.758$, $p = .475$, and Experiment 1b, $F(1, 38) = 3.387$, $MSE = 44.988$, $p = .074$. Note that the use of a go/no-go task does not permit RT analysis for the non-target (vehicle) category.

2.2 Experiment 2

2.2.1 Behavioral data—In Experiment 2, participants performed an animal/nonanimal go/no-go categorization task while their ERPs were recorded at the same time from a whole head array of electrodes (Figure 1); an objects' scene background was either maintained or phase-randomized to examine the effect of scene context on rapid object categorization. The mean accuracy and median reaction time in each of the experimental conditions are shown in Table 2. Note that the accuracy measures for animals are correct go responses (hits), whereas the accuracy measures for vehicles are correct no-go responses (correct rejections). A two-way (object category $\times$ scene background) ANOVA performed on the mean accuracy data showed a significant main effect of object category, $F(1, 15) = 8.011$, $MSE = .005$, $p = .013$, indicating that the mean accuracy for detecting animals was higher than the mean accuracy for detecting vehicles. There was also a significant main effect of scene background, $F(1, 15) = 17.234$, $MSE = .001$, $p = .001$, demonstrating that the presence of the original scene background resulted in higher accuracy regardless of object category. There was no interaction between object category and scene background, $F(1, 15) = 3.945$, $MSE = .001$, $p = .066$.

For the analysis of the reaction time data, a paired-samples t-test was performed to examine whether scene background has an effect on detection of target items. The results showed that the median reaction time for detecting animals appearing with the original scene background was significantly faster than the median reaction time for detecting animals appearing with a phase-randomized scene background, $t(1, 15) = -2.593$, $p = .020$.

2.2.2 Electrophysiological data—The presence of an animal and the presence of a scene background, respectively, elicited effects of object category (Figure 2) and scene background (Figure 3), as was apparent in the grand-averaged ERPs along the midline of electrodes together with the corresponding difference waves. For both effects, a frontal as well as a parietal positivity were elicited by the presence of an animal and by the presence of a scene background. These positivities began during the negative-going N2 deflection, which could be described as an attenuation of the frontal negativity for trials containing animals (Figure 2) or a comparable attenuation of the frontal negativity for trials accompanied by a scene background (Figure 3). A long-lasting centro-parietal positivity, the late positive potential, ensued in response to trials containing an animal (Figure 2) as well as for trials accompanied by a scene background (Figure 3). A similar extent of the object category effect was also seen in the original and phase-randomized background conditions;

importantly, analysis of Animal/Vehicle difference waves showed that there were differences in the spatiotemporal character of this object category effect (Figure 4).

2.2.2.1. Object Category Effect: As depicted in Figure 2, cluster permutation testing of the difference between ERPs elicited by animal and vehicle pictures revealed this object category effect to be significant, as took the form of the one significant positive cluster, $t\text{-sum} = 57663.47$, $p < .001$. This cluster consisted of a bilateral frontal positivity occurring 180–320 ms post-stimulus onset that progressed to a bilateral posterior positivity occurring between 320 and 700 ms. This one significant cluster indicated that the frontal negativity was more negative in trials containing vehicles and that there was an overlapping posterior late positive potential for trials containing animals. That is, the presentation of an animal elicited an attenuation of the frontal negativity (i.e., a positivity) and an increase in the late positive potential, which together emerged not as separate clusters but as one significant positive cluster. There were no negative clusters.

2.2.2.2. Scene background effect: As depicted in Figure 3, cluster permutation testing of the difference between ERPs elicited by objects with and without their original scene background revealed a significant scene background effect, as took the form of the one significant positive cluster, $t\text{-sum} = 25634.07$, $p < .001$. This cluster consisted of a bilateral frontal positivity occurring from 220–300 ms that was superimposed onto a bilateral late posterior positivity apparent from 180 until 540 ms post-stimulus onset. This one significant cluster indicated that the frontal negativity was larger in trials with a phase-randomized background, whereas the late positive potential was larger in trials with an original scene background. Three negative clusters were not significant, $p > .225$, as were two other positive clusters, $p > .212$.

2.2.2.3. Object category $\times$ scene background difference: As depicted in the ERPs and difference waves of Figure 4, there was a numerical increase in the parietal positivity produced by the presence of an animal in trials where the original background was present during a period 200–300 ms post-stimulus onset. However, when assessing the significance between the two relevant difference waves depicted in Figure 4, the positive cluster was not significant, $t\text{-sum} = 1778.18$, $p = .094$. There were two other positive clusters, $p > .212$, and two negative clusters, $p > .225$, none of which were significant. That none of these clusters were significant does not rule out the possibility of differences in the spatiotemporal character of the object category cluster as a function of background condition, to which the next section turns.

2.2.2.4. Object category effect with the original and phase-randomized scene background: To investigate the effect of scene backgrounds on early object processing, analyses focused upon a comparison of the onset latency of the object category effect when the original background was present against the corresponding onset latency of the object category effect when the background was phase-randomized. That is, the onset latencies of the object category effect in the different scene background conditions were investigated: (1) the object category effect with original backgrounds (i.e., “animals with original backgrounds” - “vehicles with original backgrounds”, and (2) the object category effect with phase-randomized backgrounds (i.e., “animals with phase-randomized backgrounds” - “vehicles with phase-randomized backgrounds”).

As depicted within the dashed box in Figure 4, in the original scene background condition, the object category cluster began at multiple anterior sites, attaining significance for each sample throughout an early time bin 120–140 ms after the onset of stimulus presentation. By contrast, in the phase-randomized scene background condition, the object category effect at anterior sites began as part of a positive cluster somewhat later, during the 140–160 ms time

bin (Figure 4, dotted box). Whether the original scene background was presented, $t\text{-sum} = 48590.72$, $p < .001$, or a phase-randomized background was presented, $t\text{-sum} = 448871.13$, $p < .001$, cluster analyses revealed that a significant positivity was elicited by the presence of an animal relative to a vehicle. There were no other significant clusters, $p > .282$. The inclusion of sample-specific t-tests in the significant object category positive cluster began earlier when the original scene background was used, rather than the phase-randomized background. That is, the object category effect began early when the original background was present. The influence of scene background information on the object category effect is visible in the difference waves at the electrodes in the frontal site, as is depicted by the vertical lines upon the upper panels of the second column of Figure 4; the first dashed line depicting the start of the 120–140 ms time bin, when the object category effect became significant with an original background, the second dotted line depicting the start of the later 140–160 ms time bin when the object category effect became significant with a phase-randomized background. As seen in the maps from Figure 4, comparison of sample specific t-tests that were included in the object category positive cluster throughout the period 200–300 ms post-stimulus onset revealed that parieto-occipital sites were only part of a significant cluster when a scene background was present.

That is, while the overall object category effect did not increase when a scene background was present (section 2.2.2.3), the object category effect began earlier at frontal sites. With the original scene background, the object category effect then showed an earlier inclusion of parieto-occipital sites in the significant positivity cluster for this effect during a period 200–300 ms post-stimulus onset. Together, this shift in the spatiotemporal character of the object category positivity cluster in the original scene background condition is thus understood to be a shift in latency of the cluster rather than an overall amplitude augmentation of the object category effect by the presence of a meaningful scene background.

2.2.2.5. Auxiliary analyses of amplitudes at single electrodes: The positivities depicted as clusters in Figures 2–4 exhibited a different distribution during the early frontal negativity (100–300 ms) from that seen during the late positive potential (300–650 ms). In an auxiliary analysis, it was thus assessed if effects were significant during these different time ranges at Fz and Pz, respectively, testing whether there was an object category or scene background effect at either latency and whether the object category effect varied as a function of scene background during either latency. As reflected by the mean ERP amplitudes and the standard error of the mean (s.e.m.), the object category effect was significant at Fz during the frontal negativity: Animal, $-6.40 \mu\text{V}$, s.e.m., $1.27 > \text{Vehicle}$, $-8.37 \mu\text{V}$, s.e.m., 1.28 , $t(15) = 4.90$, $p = .00019$; and at Pz during the late positive potential: Animal, $8.88 \mu\text{V}$, s.e.m., $1.40 > \text{Vehicle}$, $2.46 \mu\text{V}$, s.e.m., 1.04 , $t(15) = 7.77$, $p = .000001$. The scene background effect was significant during the frontal negativity: Original Scene Background, $-8.28 \mu\text{V}$, s.e.m., $1.34 > \text{Phase-randomized Scene Background}$, $-9.66 \mu\text{V}$, s.e.m., 1.34 , $t(15) = 7.72$, $p = .000001$; and at Pz during the late positive potential: Original Scene Background, $5.11 \mu\text{V}$, s.e.m., $1.16 > \text{Phase-randomized Scene Background}$, $3.03 \mu\text{V}$, s.e.m., 1.08 , $t(15) = 5.76$, $p = .00004$. The object category effect was neither significantly stronger with a background during the frontal negativity: object category effect with original scene backgrounds, $2.06 \mu\text{V}$, s.e.m., $0.53 \approx \text{object category effect with phase-randomized scene backgrounds}$, $1.89 \mu\text{V}$, s.e.m., 0.40 , $t(15) = 0.34$, $p = .739$; nor during the late positive potential: object category effect with original scene backgrounds, $6.30 \mu\text{V}$, s.e.m., $0.84 > \text{object category effect with phase-randomized backgrounds}$, $6.54 \mu\text{V}$, s.e.m., 0.91 , $t(15) = -0.39$, $p = .701$. That is, effects of object category and scene background upon ERP amplitude were significant and additive during both the frontal negativity and late positive potential latency ranges. The pattern of significance as a function of object category and scene background was thus identical for the frontal negativity and the late positive potential.

3. Discussion

The primary goal of the current study was to investigate how scene context affects rapid object recognition. In Experiment 1, participants were asked to perform an animal/nonanimal go/no-go categorization task in which they had to make a response each time a briefly presented picture contained an animal and withhold their response otherwise. Additionally, the effect of scene context was manipulated by maintaining, deleting, or phase-randomizing an object's scene background. The results showed that participants were able to perform the task accurately and quickly. Moreover, participants' accuracy and reaction times were significantly better for target objects appearing in their original scene backgrounds, demonstrating that scene context facilitates object recognition even when an image is presented briefly.

In Experiment 2, the effect of scene context on rapid object recognition was further examined by inspecting both behavioral and electrophysiological responses. Participants performed an animal/nonanimal go/no-go categorization task while their ERPs were recorded at the same time. The effect of scene context was manipulated by either maintaining or phase-randomizing an object's background information. The results showed that participants responded more accurately and quickly to objects embedded in their original scene backgrounds, confirming the importance of scene context for object recognition. In addition, compared to the original scene background condition, the onset latency of the differential activity between animal and vehicle items in the frontal electrode sites was delayed by about 20 ms in the phase-randomized scene background condition, thus suggesting that scene context facilitates object processing in the visual system. Moreover, animals or vehicles appearing in phase-randomized scene backgrounds elicited larger frontal negativities and smaller late positive potentials, suggesting that the lack of scene context makes object recognition a more demanding process, and object processing is less efficient when scene context is eliminated (see below).

The results of the present study are consistent with previous research showing that object categorization in natural scenes can be achieved efficiently (e.g., Bacon-Macé, Macé, Fabre-Thorpe, & Thorpe, 2005; Keysers & Perrett, 2002; Macé, Joubert, Nespoulous, & Fabre-Thorpe, 2009; Thorpe et al., 1996). In the current study, participants were able to perform a rapid visual superordinate categorization task (animals vs. vehicles) accurately and quickly, demonstrating that a large amount of information can be extracted from a briefly presented scene and mediates such remarkable categorization performance. The current study further reveals that scene context is capable of modulating object recognition even when scene background information is presented simultaneously with an object for a short time. The results of Experiments 1 and 2 consistently showed that participants' accuracy and reaction times were better for objects appearing with their original scene backgrounds. Therefore, the current results support previous studies showing that consistent scene background benefits object recognition (Biederman, Kosslyn, & Osherson, 1995; Boyce & Pollatsek, 1992; Boyce et al., 1989; Davenport, 2007; Davenport & Potter, 2004; Joubert et al., 2008; Palmer, 1975).

The results of the current study also support the hypothesis that scene and object information might be processed in parallel with similar temporal dynamics of visual processing and interact with each other very early in the visual pathway (e.g., Joubert et al., 2008; Joubert, Rousselet, Fize, & Fabre-Thorpe, 2007). The results of Experiment 2 showed that the onset latency of the frontal go/no-go ERP difference, which is considered an index of the minimum visual processing needed to differentiate a target from a distractor (e.g., Delorme, Rousselet, Macé, & Fabre-Thorpe, 2004; Goffaux, Jacques, Mouraux, Oliva, Schyns, & Rossion, 2005; Schmitt, Münte, & Kutas, 2000; Thorpe et al., 1996; VanRullen & Thorpe,

2001), occurred as early as 120 ms after stimulus presentation for objects embedded in their original scene backgrounds. However, the frontal differential activity between animal and vehicle pictures became significant at a later time, 140 ms, in the phase-randomized scene background condition. The results therefore suggest that scene context reduces the time needed for object categorization, given that the onset latency of frontal differential activity was delayed by about 20 ms when objects were presented without a meaningful scene background.

It is worth noting that in the current study, the onset latency of the frontal difference wave between animal and vehicle items is shorter (e.g., 120 ms in the original scene background condition) than in previous studies on object categorization in natural scenes (e.g., 150 ms in Thorpe, et al., 1996). Such latency differences may be due to subtle changes in stimulus properties. In our study, for example, the image size was enlarged to $8.2^\circ \times 13.7^\circ$ of visual angle 5° in Thorpe et al., 1996). Given that research has shown that stimulus size is (instead of $5^\circ \times$ capable of modulating the latency of some ERP components (e.g., longer latencies of P1 for smaller stimuli in Busch, Debener, Kranczioch, Engel, & Herrmann, 2004), it is possible that the larger images used in the current study increase the amount of energy in the images and therefore reduce the latency of the differential activity between object categories. The issue of how stimulus size or other factors may affect the onset latency of the differential activity between various object categories deserves future investigation.

The effect of scene context on the amplitudes of the frontal negativity confirms that object recognition becomes a more demanding task when scene context is eliminated. The results of Experiment 2 showed that the ERPs elicited by vehicle pictures were more negative-going in an early time window at the frontal electrode sites; the results were consistent with previous studies which showed an enhancement in the frontal negativity for no-go (e.g., vehicles) as compared with go (e.g., animals) stimuli. The current results also showed that scene context modulated the amplitude of the frontal negativity in different scene background conditions. That is, objects appearing without their original scene backgrounds elicited larger frontal negativities regardless of their category. Given that the enhanced frontal negativity is widely recognized as an index of increased inhibitory processing (e.g., Bokura et al., 2001; Eimer, 1993; Falkenstein et al., 1999), the current result suggests that greater effort (or need for response inhibition) is required for processing objects embedded in random and meaningless scene backgrounds (e.g., phase-randomized backgrounds). Therefore, the amplitude of the frontal negativity is enhanced because the lack of scene context makes object recognition a more demanding process, and participants need to devote more effort to withholding their responses accordingly.

The effect of scene context on the amplitudes of the late positive potential further suggests that object processing is more efficient when meaningful scene background information is presented. The results of Experiment 2 showed that the ERPs elicited by animal pictures were more positive-going in a later time window at the parietal electrode sites. This finding is consistent with previous studies showing that target stimuli elicit larger late positive potentials than nontarget stimuli in an object categorization task, suggesting that the enhanced amplitude of the late positive potential reflects increased attentional resources devoted to target events (Codispoti et al., 2006; Duncan-Johnson & Donchin, 1982; Ferrari et al., 2008; Friedman, Simson, Ritter, & Rapin, 1975; Kok, 1997, 2001). However, the amplitude of the late positive potential is reduced when participants are asked to process objects appearing with phase-randomized scene backgrounds, suggesting that fewer attentional resources are available in these conditions. Therefore, contextual information may modulate the allocation of attention to objects embedded in scenes, and thus affect the process of object recognition.

These results are therefore consistent with studies suggesting that the holistic scene representation provides an efficient way for priming typical objects and their locations in a scene and thus facilitate object processing (Bar, 2004; Bar et al., 2006; Torralba, 2003; Torralba, Oliva, Castelhano, & Henderson, 2006; Torralba & Sinha, 2001). Torralba et al. (2006), for example, demonstrated that scene context information is available rapidly enough to affect attentional allocation during the first fixation on a scene by constraining the potential locations of a target and directing attention to the most probable target location (e.g., by searching for a painting on the wall). Therefore, global scene processing can be used to constrain local feature analysis and enhance object recognition in natural scenes (Oliva & Torralba, 2006). As a result, deleting scene context impairs object categorization performance.

Taken together, the present study has shown that a considerable amount of information regarding an object and its scene context can be extracted rapidly. Moreover, the rapidly extracted scene representation can be used to modulate object categorization in an early stage of visual processing. Consequently, processing an object appearing with its original scene background is more efficient than processing an object appearing with a meaningless background.

4. Experimental procedure

4.1 Experiment 1

4.1.1 Participants—Thirty-Eight undergraduates (18 in Experiment 1a: ten males, ages 18–23 years, mean age 19.7 years; 20 in Experiment 1b: nine males, ages 18–25 years, mean age 19.9 years) from North Dakota State University participated in the study for course credit. Informed written consent was obtained from the participants. All the participants self-reported that they had normal or corrected-to-normal vision. In addition, they were naïve to the purpose of the study. The experimental protocol was approved by the North Dakota State University Institutional Review Board for the protection of human participants in research.

4.1.2 Apparatus—The stimuli were presented centrally on a 17-inch CRT monitor with a refresh rate of 100 Hz. Responses for the experimental trials were collected through the left mouse button. The experiment was programmed using Presentation software (Neurobehavioral Systems, <http://nbs.neurobs.com/>). Participants were tested individually in a room with normal interior lighting. The viewing distance was held constant at 90 cm.

4.1.3 Stimuli—In Experiment 1a and 1b, a total of 320 natural scenes were taken from a large commercial CD-ROM library (Corel Stock Photo Library). Half the images contained a wide range of animals, such as mammals, birds, fish, insects, and reptiles. The other half of the images contained various vehicles, such as cars, trucks, trains, motorcycles, airplanes, helicopters, and boats. The size and position of these objects in a single picture were as varied as possible. Based on these 320 natural scenes, the other two sets of 320 stimuli were created by either phase-randomizing or deleting the scene background surrounding the object so that the object would appear with severely limited background information. Note that phase randomization disrupts the structure of an image but preserves the contrast energy of an image so that the low-level image properties (e.g., overall luminance, spatial frequency) are the same as in the original image. Therefore, observed differences in participants' performance cannot be attributed to changes in low-level image features. Using Matlab (The MathWorks, <http://www.mathworks.com/>), phase randomization can be accomplished in the following five steps: (1) Take an image's Fourier transform, (2) calculate the phase and amplitude at each frequency, (3) add random noise to the phase

information at each frequency, (4) re-combine the amplitude information with the new phase information, and (5) perform an inverse Fourier transform on the result of step 4.

Due to these manipulations, three different versions of each picture were used in Experiment 1a: objects appeared in their original scene background, in a blank gray background, and in a phase-randomized background. Each participant saw each object three times: once in each background condition. The stimuli were randomly presented during the experiment; moreover, an object was not repeated in the same experimental block. The stimuli used in Experiment 1b were identical to Experiment 1a except that a Gaussian blur was applied to smooth the edges of the box that surrounding an object in the blank and phase-randomized background conditions. During the experiment, each image subtended $8.2^{\circ} \times 13.7^{\circ}$ of visual angle on the computer screen. Figure 5 shows sample stimuli and manipulations for different conditions in Experiments 1a and 1b.

4.1.4 Procedure and Design—Figure 6 illustrates the sequence of events in each trial of Experiments 1a and 1b. A trial began with a central cross as a fixation point for a 600–900 ms random duration. Then a picture was briefly displayed at the center of the screen for 20 ms. Participants had to press the left mouse button as quickly and as accurately as possible if the picture contained an animal (go response) and to withhold a response otherwise (no-go response). The next trial started 1000 ms after the button-press response was made or from the end of the 1000 ms response window of the picture on no-go trials. Only responses to the go stimuli within 1 s were regarded as correct. Longer reaction times were considered as no-go responses.

Experiments 1a and 1b used a 2 (object category: animal vs. vehicle) $\times$ 3 (scene background: original, blank, phase-randomized) factorial design. Participants first completed a practice session of 24 trials, 4 in each of the 6 conditions. Following the 24 practice trials, participants completed 960 experimental trials, 160 in each of the 6 conditions. The 960 experimental trials were presented randomly for each participant, in eight blocks of 120 trials each.

4.2 Experiment 2

4.2.1 Participants—Sixteen undergraduates (thirteen males, ages 18–21 years, mean age 19.1 years) from North Dakota State University participated in the study for course credit. All the participants self-reported that they had normal or corrected-to-normal vision. Additionally, they had not participated in the previous study. All participants completed a health survey and reported no history of neurological problems, serious head-injuries, or serious psychiatric conditions. In addition, informed written consent was obtained from the participants. The experimental protocol was approved by the North Dakota State University Institutional Review Board for the protection of human participants in research.

4.2.2 Stimuli—Experiment 2 was identical to Experiment 1, except that only the set of 320 objects with phase-randomized scene backgrounds were used to examine the effect of deleting scene backgrounds on rapid object categorization. Therefore, each participant saw each object twice: with a normal and a phase-randomized scene background, respectively. During the experiment, the stimuli were randomly presented; moreover, an object was not repeated in the same experimental block.

4.2.3 Procedure and Design—Experiment 2 was identical to Experiment 1, except that participants' responses for the experimental trials were collected through a response button that was held in their dominant hands. Participants were asked to press the response button

as quickly and as accurately as possible if the picture contained an animal, and to withhold a response otherwise.

Experiment 2 used a 2 (object category: animal vs. vehicle) $\times$ 2 (scene background: original vs. phase-randomized) factorial design. Participants first completed a practice session of 32 trials, 8 in each of the 4 conditions. Following the 32 practice trials, participants completed 640 experimental trials, 160 in each of the 4 conditions. The 640 experimental trials were presented randomly for each participant, in eight blocks of 80 trials each.

4.2.4 Behavioral Recording and Analysis—Presentation software (Neurobehavioral Systems, <http://nbs.neurobs.com/>) was used to present the stimuli and record the behavioral responses. The stimuli were presented centrally on a 17-inch CRT monitor with a refresh rate of 100 Hz; responses for the experimental trials were collected through a response button. Participants were tested individually in a dimly lit, sound attenuated, electrically shielded room.

For the behavioral data analysis, the proportion of correct responses for different conditions was calculated and submitted to a repeated-measures ANOVA. The median reaction times for animals presented with and without their original scene backgrounds were calculated and submitted to a paired-samples t-test to examine whether there was a difference in response in between two given conditions.

4.2.5 Evoked-potential Recording and Analysis—The electroencephalogram (EEG) was recorded from 64 scalp sites with an Active Two BioSemi electric system (<http://www.biosemi.com>; BioSemi, Amsterdam, The Netherlands) and the electrooculogram (EOG) was recorded from six electrodes located at the outer canthi and above and beneath each eye. The electrode offset was kept below 25 mV. The EEG sampling rate was 512 Hz with a pass-band from DC to 150 Hz. In lieu of the “ground” electrode used by conventional systems, Biosemi uses two separate electrodes: the “Common Mode Sense” active electrode and the “Driven-Right-Leg” passive electrode. Further information on reference and grounding conventions can be found at <http://www.biosemi.com/faq/cms&drl.htm>. The data were re-referenced to the average of both mastoids off-line.

The data were analyzed using BESA 5.1.8 (Brain Electric Source Analysis, Gräefeling, Germany). Automated artifact rejection criteria of $\pm 120 \mu\text{V}$ were applied between -100 and +700 ms before averaging to discard trials during which an eye movement, a blink, or amplifier blocking had occurred. Only the remaining trials with correct responses were averaged for each condition and for each participant. ERPs were then averaged for each electrode over all 16 participants, and four datasets retained: (1) animals with original backgrounds, (2) vehicles with original backgrounds, (3) animals with phase-randomized backgrounds, and (4) vehicles with phase-randomized backgrounds. Baseline correction was performed using the 100 ms of pre-stimulus activity. The data were also low-pass filtered at 35 Hz before analysis. Baseline-corrected ERPs at each electrode over all 16 participants were then collapsed across object category, yielding two additional datasets: (5) original background and (6) phase-randomized background. ERPs were also collapsed across scene background, yielding another two data sets: (7) animal and (8) vehicle. Five sets of difference waves were computed by the subtraction of datasets, the first four of which were: (1) object category effect (e.g., “animal” - “vehicle”), (2) scene background effect (e.g., “original background” - “phase-randomized background”), (3) object category effect with original backgrounds (e.g., “animals with original backgrounds” - “vehicles with original backgrounds”), and (4) object category effect with phase-randomized backgrounds (e.g., “animals with phase-randomized backgrounds” - “vehicles with phase-randomized

backgrounds”). To determine if the object category effect varied as a function of scene background condition, this object category effect with phase-randomized backgrounds difference wave was subtracted from the object category effect with original backgrounds difference wave, as yielded the difference of these difference waves, the fifth object category $\times$ scene background difference wave.

Five cluster permutation tests were used to assess the significance of each of these differences in a manner that solves the multiple comparison problem with a nonparametric statistical test (Maris & Oostenveld, 2007; Maris, 2004; Haegens, Osipova, Oostenveld, & Jensen, 2010) using the Matlab-based FieldTrip toolbox (<http://fieldtrip.fcdonders.nl/>). The approach was that a randomization distribution statistic is constructed and used to evaluate statistically significant differences between conditions. That is, for each channel at each sample post-stimulus onset, dependent-sample *t* statistics were computed and an algorithm formed spatiotemporal patterns of differences (clusters), which were based upon these sample-specific *t*-tests. Criteria for membership in a cluster were that sample-specific *t*-tests at a given time in a given channel had 10 neighboring electrodes that simultaneously exhibited *t*-values that exceeded the corresponding two-tailed univariate *t*-test in one channel with critical α set to .05. These criteria thus excluded implausibly focal false alarm differences from being included in clusters as well as precluding bridges from being formed between any genuinely distinct clusters. For each cluster, the maximum of the sum of the sample-specific *t* statistics was then used to determine a cluster-level statistic “*t-sum*”, which was then used to test the overall significance of that cluster. This cluster-level statistic was evaluated by randomizing the data across the two conditions and recalculating that test statistic 1000 times to construct a reference distribution. The cluster-level statistic from the actual data was compared this reference distribution with a two-tailed critical α set to .025 as appropriate.

Grand average ERPs were over-plotted for the midline alongside the corresponding difference wave, and maps of the polarity of the difference integrated across time bins, highlighting upon each map electrodes that were members of a significant cluster in the difference wave throughout that time bin.

Cluster analyses revealed one significant positive cluster for the object category effect and the scene background condition, though the scalp distribution of the effect varied as a function of time. To test the significance of effects during a time window at single electrodes appropriate to the frontal negativity and the late positive potential, sections of the original 512-Hz digitized individual difference waves were resampled at 10 kHz and 100 ms windows of integration were centered upon the peak of each individual difference wave at the relevant electrode. ERPs and difference waves of interest were resampled using cubic spline interpolation (de Boor, 1978) at 10 kHz. To calculate the resampled waveform within the measurement windows, original samples were used between and inclusive of the nearest sample before the onset of the window up until the nearest sample after the whole duration of the window. A functionally-identical technique for quantification of amplitudes has been contributed to the Matlab-based EEGLAB toolbox (Delorme & Makeig, 2004).

In the auxiliary analysis of amplitude measurements that was made with this technique for each of three effects separately. These three effects were: (1) object category effect (“animal” vs. “vehicle”), (2) scene background effect (“original background” vs. “phase-randomized background”), and (3) the object category $\times$ scene background difference, which was a difference of two difference waves (i.e., the “object category effect with original backgrounds” difference wave minus the “object category effect with phase-randomized backgrounds” difference wave). For the data from individual participants, the amplitude of these relevant waveforms at Fz was derived for a 100 ms window of integration centered

upon the individual peak of the relevant difference wave at Fz in between 100 and 300 ms post-stimulus onset, as well as the amplitude of the waveforms at Pz during the late positive potential during a time window centered upon the individual peak of the relevant difference wave at Pz in between 300 and 650 ms. The choice of electrodes and the time windows where individual difference wave peaks were identified were based upon the timing and distribution of the effects evident in grand-averages in Figures 2–3. This approach of centering windows on individual peaks of relevant waveforms was adapted from Campbell, Winkler, and Kujala (2007), with a view to enhancing the accuracy of amplitude measurement to improve sensitivity so as to identify any qualitative difference in the pattern of significant differences for the frontal negativity and the late positive potential. The relevant differences were: (1) for the object category effect (i.e., “animal” - “vehicle”), frontal negativity peaks were in a range from 117.23 to 269.00 ms, mean latency 212.64 ms, s.e.m. 12.10, $n = 16$; late positive potential peaks were in a range from 326.55 to 620.43 ms, mean latency 460.30 ms, s.e.m. 20.49, $n = 16$, (2) for the scene background effect (i.e. “original background” - “phase-randomized background”), frontal negativity peaks were in a range from 125.95 to 297.02 ms, mean latency 253.23 ms, s.e.m. 10.21; late positive potential peaks were in a range from 307.78 to 603.61 ms, mean latency 461.04 ms, s.e.m. 23.25, $n = 16$, and (3) for the object category $\times$ scene background difference (i.e., “object category effect with the original background” - “object category effect with the phase-randomized background”): frontal negativity peaks were in a range from 100.00 to 300.00 ms, mean latency 190.87 ms, s.e.m. 19.08; late positive potential peaks were in a range from 300.00 to 649.00 ms, mean latency 472.51 ms, s.e.m. 29.65, $n = 16$.

For each of these three effects, a t-test was then used to ascertain the significance of the difference between waveforms at each electrode in the respective time window, as meant a total of six dependent samples t-tests within this auxiliary analysis, each of which employed a two-tailed critical α set to .05.

Acknowledgments

This project was supported by Grant Number 1P20 RR020151 from the National Center for Research Resources (NCRR), a component of the National Institutes of Health (NIH). The project was also supported by the National Science Foundation under Grant Number BCS-0443998. We thank the reviewers for their helpful comments on the manuscript; we also thank Tom Campbell for his help with the ERP data analysis. Finally, we thank Olga Sysoeva and Eric Maris for their statistical suggestions.

References

- Bacon-Macé, Ng; Macé, MJM.; Fabre-Thorpe, Ml; Thorpe, SJ. The time course of visual processing: Backward masking and natural scene categorisation. *Vision Research*. 2005; 45:1459–1469. [PubMed: 15743615]
- Bar M. Visual objects in context. *Nature Reviews Neuroscience*. 2004; 5:617–629.
- Bar M, Kassam KS, Ghuman AS, Boshyan J, Schmidt AM, Dale AM, et al. Top-down facilitation of visual recognition. *Proceedings of the National Academy of Science*. 2006; 103:449–454.
- Biederman, I.; Kosslyn, SM.; Osherson, DN. *Visual cognition: An invitation to cognitive science*. 2. Vol. 2. Cambridge, MA US: The MIT Press; 1995. Visual object recognition; p. 121-165.
- Biederman I, Mezzanotte RJ, Rabinowitz JC. Scene perception: Detecting and judging objects undergoing relational violations. *Cognitive Psychology*. 1982; 14:143–177. [PubMed: 7083801]
- Bokura H, Yamaguchi S, Kobayashi S. Electrophysiological correlates for response inhibition in a Go/NoGo task. *Clinical Neurophysiology*. 2001; 112:2224–2232. [PubMed: 11738192]
- Boyce SJ, Pollatsek A. Identification of objects in scenes: The role of scene background in object naming. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 1992; 18:531–543.

- Boyce SJ, Pollatsek A, Rayner K. Effect of background information on object identification. *Journal of Experimental Psychology: Human Perception and Performance*. 1989; 15:556–566. [PubMed: 2527962]
- Boyce SJ, Pollatsek A, Rayner K. Effect of background information on object identification. *Journal of Experimental Psychology: Human Perception and Performance*. 1989; 15:556–566. [PubMed: 2527962]
- Busch NA, Debener S, Kranczioch C, Engel AK, Herrmann CS. Size matters: Effects of stimulus size, duration and eccentricity on the visual gamma-band response. *Clinical Neurophysiology*. 2004; 115:1810–1820. [PubMed: 15261860]
- Campbell TA, Winkler I, Kujala T. N1 and the mismatch negativity are spatiotemporally distinct ERP components: Disruption of immediate memory by auditory distraction can be related to N1. *Psychophysiology*. 2007; 44:530–540. [PubMed: 17532805]
- Codispoti M, Ferrari V, Junghöfer M, Schupp HT. The categorization of natural scenes: brain attention networks revealed by dense sensor ERPs. *Neuroimage*. 2006; 32:583–591. [PubMed: 16750397]
- Davenport JL. Consistency effects between objects in scenes. *Memory & Cognition*. 2007; 35:393–401.
- Davenport JL, Potter MC. Scene Consistency in Object and Background Perception. *Psychological Science*. 2004; 15:559–564. [PubMed: 15271002]
- de Boor, C. A practical guide to splines. New York: Springer-Verlag; 1978.
- de Gelder B, Meeren HKM, Righart R, Van den Stock J, van de Riet WAC, Tamietto M. Beyond the face: Exploring rapid influences of context on face processing. *Progress in Brain Research*. 2006; 155:37–48. [PubMed: 17027378]
- Delorme A, Makeig S. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics. *Journal of Neuroscience Methods*. 2004; 134:9–21. [PubMed: 15102499]
- Delorme A, Rousselet GA, Macé MJM, Fabre-Thorpe M. Interaction of top-down and bottom-up processing in the fast visual analysis of natural scenes. *Cognitive Brain Research*. 2004; 19:103–113. [PubMed: 15019707]
- Duncan-Johnson CC, Donchin E. The P300 component of the event-related brain potential as an index of information processing. *Biological Psychology*. 1982; 14:1–52. [PubMed: 6809064]
- Eimer M. Effects of attention and stimulus probability on ERPs in a Go/Nogo task. *Biology Psychology*. 1993; 35:123–138.
- Falkenstein M, Hoormann J, Hohnsbein J. ERP components in Go/Nogo tasks and their relation to inhibition. *Acta Psychologica*. 1999; 101:267–291. [PubMed: 10344188]
- Ferrari V, Codispoti M, Cardinale R, Bradley MM. Directed and motivated attention during processing of natural scenes. *Journal of Cognitive Neuroscience*. 2008; 20:1753–1761. [PubMed: 18370595]
- Friedman D, Simson R, Ritter W, Rapin I. The late positive component (P300) and information processing in sentences. *Electroencephalography and Clinical Neurophysiology*. 1975; 38:255–262. [PubMed: 46803]
- Ganis G, Kutas M. An electrophysiological study of scene effects on object identification. *Cognitive Brain Research*. 2003; 16:123–144. [PubMed: 12668221]
- Goffaux, V.; Jacques, C.; Mouraux, A.; Oliva, A.; Schyns, PG.; Rossion, B. Diagnostic colours contribute to the early stages of scene categorization: Behavioural and neurophysiological evidence. *Visual Cognition*. 2005; 12:878–892.
- Gordon RD. Attentional allocation during the perception of scenes. *Journal of Experimental Psychology: Human Perception and Performance*. 2004; 30:760–777. [PubMed: 15301623]
- Gordon RD. Selective attention during scene perception: Evidence from negative priming. *Memory & Cognition*. 2006; 34:1484–1494.
- Haegens S, Osipova D, Oostenveld R, Jensen O. Somatosensory working memory performance in humans depends on both engagement and disengagement of regions in a distributed network. *Human Brain Mapping*. 2010; 31:26–35. [PubMed: 19569072]
- Intraub H. Rapid conceptual identification of sequentially presented pictures. *Journal of Experimental Psychology: Human Perception and Performance*. 1981; 7:604–610.

- Joubert OR, Fize D, Rousselet GA, Fabre-Thorpe MI. Early interference of context congruence on object processing in rapid visual categorization of natural scenes. *Journal of Vision*. 2008; 8:1–18. [PubMed: 19146341]
- Joubert OR, Rousselet GA, Fize D, Fabre-Thorpe MI. Processing scene context: Fast categorization and object interference. *Vision Research*. 2007; 47:3286–3297. [PubMed: 17967472]
- Keysers C, Perrett DI. Visual masking and RSVP reveal neural competition. *Trends in Cognitive Sciences*. 2002; 6:120–125. [PubMed: 11861189]
- Kok A. Event-related-potential (ERP) reflections of mental resources: a review and synthesis. *Biological Psychology*. 1997; 45:19–56. [PubMed: 9083643]
- Kok A. On the utility of P3 amplitude as a measure of processing capacity. *Psychophysiology*. 2001; 38:557–577. [PubMed: 11352145]
- Kutas M, Hillyard SA. Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science*. 1980; 207:203–205. [PubMed: 7350657]
- Macé MJM, Joubert OR, Nespoulous JL, Fabre-Thorpe M. The time-course of visual categorizations: You spot the animal faster than the bird. *PLoS ONE*. 2009; 4:e5927. [PubMed: 19536292]
- Maris E. Randomization tests for ERP-topographies and whole spatiotemporal data matrices. *Psychophysiology*. 2004; 41:42–151.
- Maris E, Oostenveld R. Nonparametric statistical testing of EEG- and MEG-data. *Journal of Neuroscience Methods*. 2007; 164:177–190. [PubMed: 17517438]
- Mudrik L, Lamy D, Deouell LY. ERP evidence for context congruity effects during simultaneous object-scene processing. *Neuropsychologia*. 2010; 48:507–517. [PubMed: 19837103]
- Oliva A, Torralba A. Building the gist of a scene: The role of global image features in recognition. *Progress in Brain Research*. 2006; 155:23–36. [PubMed: 17027377]
- Palmer SE. The effects of contextual scenes on the identification of objects. *Memory & Cognition*. 1975; 3:519–526.
- Potter MC. Meaning in visual search. *Science*. 1975; 187:965–966. [PubMed: 1145183]
- Potter MC. Short-term conceptual memory for pictures. *Journal of Experimental Psychology: Human Learning and Memory*. 1976; 2:509–522. [PubMed: 1003124]
- Righart R, de Gelder B. Context influences early perceptual analysis of faces: An Electrophysiological study. *Cerebral Cortex*. 2006; 16:1249–1257. [PubMed: 16306325]
- Rousselet GA, Fabre-Thorpe MI, Thorpe SJ. Parallel processing in high-level categorization of natural images. *Nature Neuroscience*. 2002; 5:629–630.
- Schmitt BM, Münte TF, Kutas M. Electrophysiological estimates of the time course of semantic and phonological encoding during implicit picture naming. *Psychophysiology*. 2000; 37:473–484. [PubMed: 10934906]
- Thorpe S, Fize D, Marlot C. Speed of processing in the human visual system. *Nature*. 1996; 381:520–522. [PubMed: 8632824]
- Torralba A. Contextual priming for object detection. *International Journal of Computer Vision*. 2003; 53:169–191.
- Torralba A, Oliva A, Castelano MS, Henderson JM. Contextual guidance of eye movements and attention in real-world scenes: The role of global features in object search. *Psychological Review*. 2006; 113:766–786. [PubMed: 17014302]
- Torralba A, Sinha P. Statistical context priming for object detection. *Proceeding of the International Conference on Computer Vision*. 2001; 1:763–770.
- VanRullen R, Thorpe SJ. The time course of visual processing: From early perception to decision-making. *Journal of Cognitive Neuroscience*. 2001; 13:454–461. [PubMed: 11388919]

Figure 1.
Layout of electrodes used in Experiment 2. Midline electrodes that are encircled with dark rings have grand-averaged ERPs and difference waves plotted in Figures 2–4.

Figure 2.

Grand-averaged ERPs (leftmost column) in response to stimuli containing a target (animal) or a non-target (vehicle) from a go/no-go task in Experiment 2, with the corresponding animal/vehicle difference waves (second column), together with the scalp distribution of the polarity of this difference wave as a function of time (rightmost panel); $n = 16$. Shaded areas on ERPs and difference waves denote a 100ms window of integration centered on the peak of the difference wave for the frontal negativity ($1.75 \mu V$; i.e., Animal: $-8.25 \mu V > Vehicle: -9.99 \mu V$) and the late positive potential ($5.72 \mu V$; i.e., Animal: $8.66 \mu V > Vehicle: 2.94 \mu V$). An analogous technique was applied to the individual waveforms.

Figure 3.

Grand-averaged ERPs (leftmost column) in response to stimuli with their original background or a phase-randomized background from a go/no-go task in Experiment 2, with the corresponding original/phase-randomized difference waves (second column), together with the scalp distribution of the polarity of this difference wave as a function of time (rightmost panel); $n = 16$. Shaded areas on ERPs and difference waves denote a 100ms window of integration centered on the peak of the difference wave for the frontal negativity ($1.33 \mu\text{V}$; i.e., Original: $-8.84 \mu\text{V} > \text{Phase-randomized: } -10.18 \mu\text{V}$) and the late positive potential ($1.50 \mu\text{V}$; i.e., Original: $8.66 \mu\text{V} > \text{Phase-randomized: } 5.05 \mu\text{V}$). An analogous technique was applied to the individual waveforms.

Figure 4.

Grand-averaged ERPs (leftmost column) in response to stimuli containing a target (animal) or a non-target (vehicle) from a go/no-go task as a function of whether the original scene background was maintained in Experiment 2, with the corresponding original/phase-randomized difference waves (second column), together with the scalp distribution of the polarity of this difference wave as a function of time (original background, upper right panels; phase-randomized background, lower right panels); $n = 16$. Significant cluster electrodes denoted are those that were included in a significant cluster throughout the entire time period. Maps are made with 20 ms intervals around the time of onset of the Animal-Vehicle positivity cluster that reflects a significant object category effect. The time of the beginning of the first window included in the object category positivity cluster is denoted with a vertical line upon the difference waves and by a box around the relevant map (dashed, original background: 120–140ms < dotted, phase-randomized background: 140–160ms).

Figure 5.

Sample stimuli in Experiments 1 and 2. Note that sample images a, b, c, f, g, and h were used in Experiment 1a; sample images a, d, e, f, i, and j were used in Experiment 1b; sample images a, c, f, and h were used in Experiment 2. In the experiments, all stimuli were presented in color. (a) a sample image used in the animal category with the original background condition (b) a sample image used in the animal category with a blank background condition (c) a sample image used in the animal category with a phase-randomized background condition (d) a sample image used in the animal category with a blank background condition, the Gaussian window version (e) a sample image used in the animal category with a phase-randomized background condition, the Gaussian window version (f) a sample image used in the vehicle category with the original background condition (g) a sample image used in the vehicle category with a blank gray background condition (h) a sample image used in the vehicle category with a phase-randomized background condition (i) a sample image used in the vehicle category with a blank background condition, the Gaussian window version (j) a sample image used in the vehicle category with a phase-randomized background condition, the Gaussian window version

Figure 6. The sequence of events within a trial in Experiments 1 and 2. In the experiments, all stimuli were presented in color.

Table 1

Summary of participants' categorization performance for each experimental condition of Experiments 1a (A) and 1b (B). The standard error of the mean is indicated in parenthesis. Asterisks indicate statistically significant differences (see text for details). Note: * $p < .05$; ** $p < .01$; *** $p < .001$

A.		Object Category	
Background		Animal	Vehicle
Accuracy (%)			
Original		97.5 (0.6)	85.4 (2.4)
Blank	*	96.9 (0.8)	82.0 (2.4)
Phase-randomized		96.6 (0.7)	82.6 (2.4)
Median RT (ms)			
Original		407 (10)	N/A
Blank	***	415 (9)	N/A
Phase-randomized		416 (9)	N/A
B.		Object Category	
Background		Animal	Vehicle
Accuracy (%)			
Original		97.3 (0.7)	90.3 (1.1)
Blank	*	96.3 (0.9)	87.3 (2.1)
Phase-randomized		95.6 (1.0)	86.6 (1.8)
Median RT (ms)			
Original		419 (14)	N/A
Blank	***	426 (13)	N/A
Phase-randomized		429 (13)	N/A

Table 2

Summary of participants' categorization performance for each experimental condition of Experiment 2. The standard error of the mean is indicated in parenthesis. Asterisks indicate statistically significant differences (see text for details).

Background	Object Category	
	Animal	Vehicle
Accuracy (%)		
Original	** 97.8 (0.9)	** 94.0 (1.3)
Phase-randomized	[96.4 (1.3)	[89.6 (1.8)
Median RT (ms)		
Original	* 385 (15)	N/A
Phase-randomized	[394 (15)	N/A

Note:

*
 $p < .05$;

**
 $p < .01$;

 $p < .001$