available at www.sciencedirect.com

ScienceDirect

www.elsevier.com/locate/brainres

BRAIN RESEARCH

Research Report

Contextual influences on rapid object categorization in natural scenes

Hsin-Mei Sun[a,*], Stephanie L. Simon-Dack[b], Robert D. Gordon[a], Wolfgang A. Teder[a]

[a]Department of Psychology, North Dakota State University, Fargo, ND 58105, USA
[b]Department of Psychological Science, Ball State University, Muncie, IN 47306, USA

ARTICLE INFO

Article history:
Accepted 15 April 2011
Available online 27 April 2011

Keywords:
Object categorization
Natural scene
Context effect
Event-related potentials (ERPs)

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/non-animal 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.

Published by Elsevier B.V.

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% 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 results showed that more than 70% 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-

* *Corresponding author*. Fax: +1 701 231 8426.
E-mail address: mei.sun@ndsu.edu (H.-M. Sun).

0006-8993/$ – see front matter. Published by Elsevier B.V.
doi:10.1016/j.brainres.2011.04.029

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 et al., 1982; Boyce and Pollatsek, 1992; Boyce et al., 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 trail 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 and 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 et al., 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 and 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 et al. (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/non-animal 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 and