Object Substitution Masking Interferes With Semantic Processing: Evidence From Event-Related Potentials

Jason E. Reiss and James E. Hoffman

Psychological Science 2006 17: 1015
DOI: 10.1111/j.1467-9280.2006.01820.x

Research Report

Object Substitution Masking Interferes With Semantic Processing

Evidence From Event-Related Potentials

Jason E. Reiss and James E. Hoffman

University of Delaware

ABSTRACT—*Object substitution masking (OSM) refers to impaired target identification caused by common onset, but delayed offset, of a surrounding dot mask. This effect has been hypothesized to reflect reentrant processes that result in the mask replacing the target representation. However, little is known about the depth of processing associated with masked targets in this paradigm. We investigated this issue by examining the effect of OSM on the N400 component of the event-related potential, which reflects the degree of semantic mismatch between a target and its context. Participants read a context word followed by a semantically related or unrelated target word surrounded by dots. As expected, delayed dot offset significantly reduced accuracy in identifying the target. The N400 amplitude was also diminished by OSM. These findings offer the first evidence that substitution interferes with target processing prior to semantic analysis, demonstrating an important difference between OSM and other visual phenomena, such as the attentional blink, in which semantic processing is independent of awareness.*

Object substitution masking (OSM) occurs when a target and surrounding dots are presented simultaneously and the dots persist after the target has been removed. Observers are unable to identify the target and often report that the area enclosed by the dots appears "blank." This effect is puzzling because, unlike in other forms of masking (metacontrast—Alpern, 1953; pattern integration and interruption—Turvey, 1973), the mask (i.e., OSM dots) does not need to be particularly close to the target (Di Lollo, Enns, & Rensink, 2000; Enns & Di Lollo, 1997; Jiang & Chun, 2001; Lleras & Moore, 2003; Neill, Hutchison, & Graves, 2002). In addition, other forms of masking depend critically on the relative onsets of the target and mask and cannot explain why delayed offset would promote interference by the mask (Di Lollo, Bischof, & Dixon, 1993).

Di Lollo et al. (2000) have proposed a reentrant-processing model of OSM that assumes the perception of visual objects depends on extended cycles of feedforward processes going from the retina to higher-level visual areas and feedback (or reentrant) processes that carry information from higher levels to lower levels (Lamme & Roelfsema, 2000). Initial feedforward processing results in a provisional, possibly erroneous, hypothesis about the identity of the stimulus. Because of the generally large receptive fields found in higher-level visual areas, this initial hypothesis will be based on a combination of mask and target information. Reentrant processes, however, can utilize the small receptive fields found in early visual areas to refine the initial hypothesis by segmenting targets from surrounding clutter (Lamme, Zipser, & Spekreijse, 2002). In this scheme, the target and dots are initially treated as one integrated object because of their common onset. Feedback processes attempt to verify this hypothesis by attempting to match the continuing input against the hypothesized shape. When the mask offset is delayed, a mismatch occurs because the dots are still highly visible and the target has been removed, so the initial hypothesis is discarded in favor of a new one based solely on the dots. The resulting percept retains no information about the presence of the target, so it is absent from awareness.

This model offers a valuable framework for understanding core issues of perception and awareness that concern object updating over space and time (Moore & Lleras, 2005; Woodman

Address correspondence to Jason E. Reiss, Department of Psychology, University of Delaware, Newark, DE 19716, e-mail: jreiss@udel.edu.

& Luck, 2003) and may be directly relevant to visual phenomena beyond masking (change blindness, inattentional blindness, and the attentional blink, or AB; Enns & Di Lollo, 2000). However, the existing OSM research has largely ignored a fundamental issue regarding this phenomenon: To what extent is the unreportable target processed by the visual system? The masked target may in fact be richly processed but simply fail to reach consciousness, as is evident, for example, in the AB. In the AB, two targets requiring identification are embedded in a stream of objects presented in rapid succession at the same location. Identification of the first target (T1) results in a short-lived deficit in the ability to detect or identify the second target (T2; Raymond, Shapiro, & Arnell, 1992). Like the masked target in OSM, T2 is often completely absent from conscious awareness (i.e., "blinked"). Recent evidence, however, suggests that T2 receives substantial processing, as categorical and semantic properties associated with T2 influence perception of a subsequent (T3) item (Shapiro, Driver, Ward, & Sorensen, 1997).

Electrophysiological evidence also indicates that the AB fails to prevent semantic analysis of T2 (Luck, Vogel, & Shapiro, 1996; Rolke, Heil, Streb, & Hennighausen, 2001; Vogel, Luck, & Shapiro, 1998). Vogel et al. used the N400 component of the event-related brain potential (ERP) to investigate whether a word that was blinked was nevertheless processed to a semantic level. The N400 (Kutas & Hillyard, 1980) refers to an ERP component whose amplitude depends on the degree of mismatch between a target and a semantic context. Vogel et al. reported that although the AB severely disrupted observers' ability to indicate semantic relatedness of the T2 word to a pretrial context word, unrelated T2s still elicited a robust N400 component; therefore, AB interference allows substantial semantic analysis but prevents access to conscious awareness. A similar approach based on ERPs could be used to examine the degree of processing for unreportable targets in OSM, but, to date, only one study has attempted to do this. Woodman and Luck (2003) utilized the N2pc ERP component, which reflects the deployment of attention to the left or right visual field (Luck & Hillyard, 1994). Participants had to detect a target occurring in one of two concurrently cued locations in opposite visual fields. As expected, delaying the offset of the OSM dots significantly reduced behavioral accuracy, but did not suppress the N2pc amplitude. These results indicate that although OSM interfered with the ability to report a target, it did not prevent the target from attracting attention to its location.

Clearly, at least some aspects of target processing are unaffected by the persisting masking dots. Woodman and Luck (2003) required their observers to discriminate simple shapes, so preserved information about the masked target might consist of fairly low-level physical features; however, researchers currently do not know whether high-level (e.g., word semantics) processing of targets in OSM is preserved, as it appears to be in the case of the AB. In the current study, we aimed to address this issue by examining whether OSM affects the semantic processing of a word as reflected in the amplitude of the N400 component. A failure of OSM to reduce N400 amplitude would be consistent with AB findings, suggesting that substitution occurs at a relatively late stage prior to conscious awareness but subsequent to semantic analysis. Alternatively, an OSM-based reduction in the N400 would provide evidence that, unlike certain physical target features, which are immune to OSM, semantic analysis is not.

METHOD

Subjects

Twelve native English-speaking, right-handed, neurologically normal volunteers (ages 18–32 years) were paid for their participation. All reported normal or corrected-to-normal color vision and acuity and provided informed consent. This experiment was approved by the University of Delaware Human Subjects Review Board.

Stimuli and Procedure[1]

Individuals participated in one 2-hr session, consisting of 40 practice and 368 experimental trials. Overall performance feedback was provided after practice and after every 92 experimental trials. Participants were instructed to read two words; the first word always appeared alone at the center of the display, and the second word appeared among nonword distractors in a random position cued by surrounding dots. Instructions emphasized that eye fixation should be maintained and blinks avoided until response; compliance with these instructions was monitored using a Tobii x50 50-Hz eyetracker (Tobii Technology, Stockholm, Sweden; 1.03° eye movement threshold) that transferred data to a SONY 2.86 GHz computer.

Testing was conducted in a dimly lit, electrically shielded room with a chin rest maintaining a 72-cm viewing distance. A fixation cross was presented until the participant initiated a trial sequence (see Fig. 1) with a mouse click. Context displays contained a context word (3–12 letters; $0.52^\circ \times 1.43$–6.03°) that temporarily replaced the fixation cross; whereas target displays contained the fixation cross, a single target word (3–7 letters; $0.52^\circ \times 1.43$–3.42°), and five distractor strings composed of randomly selected letters. All six strings in a target display had the same letter count; they were arranged along the vertical (two above and two below) and horizontal (one to the left and one to the right) axes of fixation. The centers of vertical strings were separated from the fixation center by 1.03° and 2.31°, respectively, and the inside edge of each horizontal string was separated from the fixation center by 1.03°. Target location (always either the top or bottom position) was cued by eight equally spaced dots forming a rectangle surrounding that word (individual dots: $0.17^\circ \times 0.17^\circ$; rectangle: $0.92^\circ \times 1.63$–3.62°, based on dot

[1]Supporting material concerning methodological details and ERP processing specifications is located on the Web at http://hoffman.psych.udel.edu/OSM/.

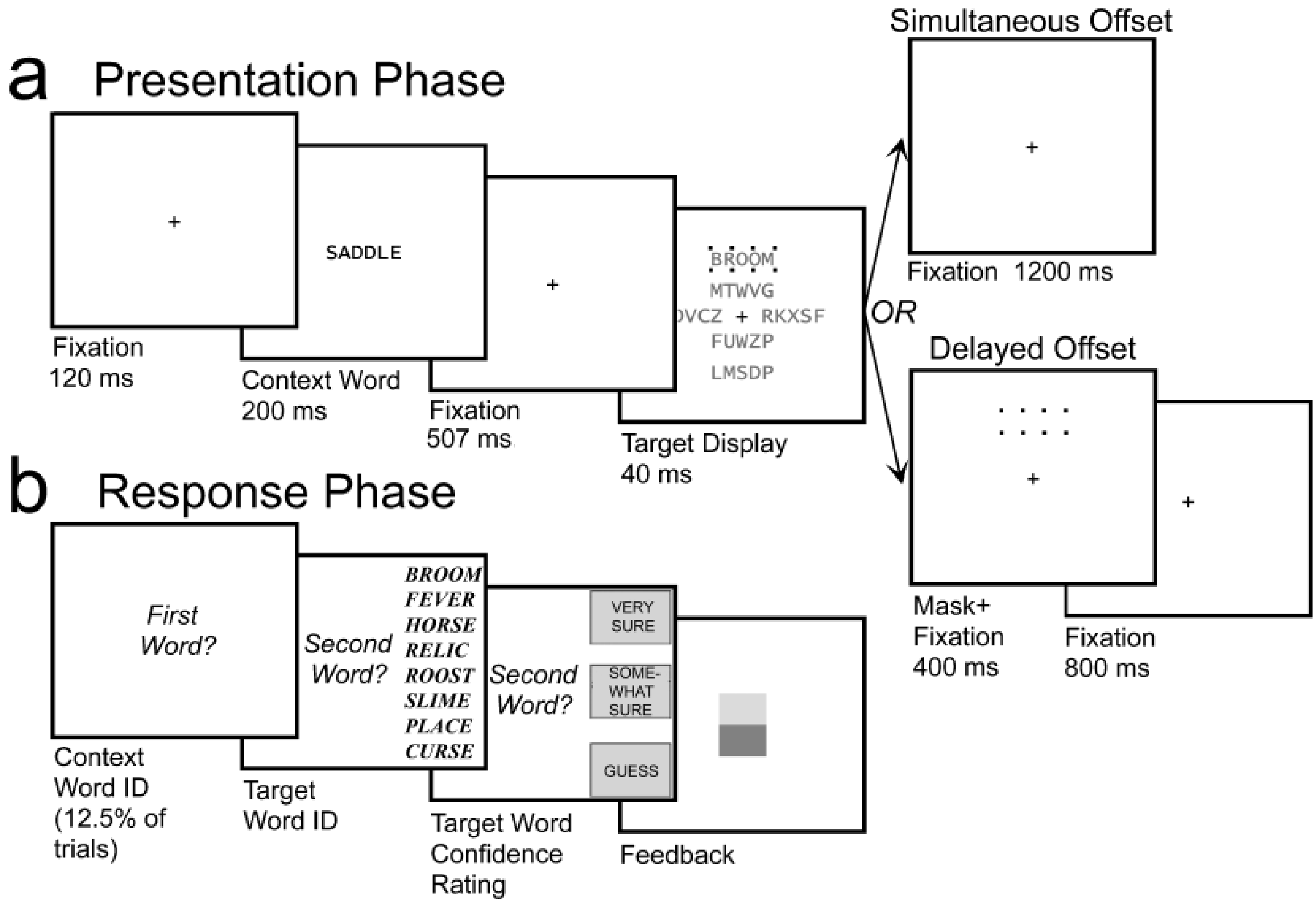

Fig. 1. Schematic diagram of the trial sequence. During the presentation phase (a), participants were shown, in order, a fixation cross, context word, fixation cross, and target display—with the target word surrounded by eight masking dots. The context and target words were either related or unrelated, and the masking dots either terminated simultaneously with the target display or were delayed in their offset (i.e., object substitution masking); a fixation-alone display concluded the sequence. The subsequent response phase (b) required participants to identify the context word on 12.5% of the trials and to identify the target word and rate their confidence in doing so on all trials. Feedback regarding response accuracy, as well as any detection of eye movements or blinks, was provided at the end of each trial. The example shown here illustrates a trial in which the context and target words were unrelated.

centers). After the target display, either the fixation cross alone was visible for 1,200 ms (simultaneous mask offset) or both the masking dots and the fixation cross were visible for 400 ms, followed by the fixation cross alone for 800 ms (delayed mask offset; OSM).

Several responses were recorded for each trial (see Fig. 1b). First, to ensure that participants read the context words, we required that they read the context word out loud on a random subset of trials (12.5%; M_{error} = 2.72% of trials tested, SE = 0.61). Second, on all trials, participants were required to identify the target word and rate their confidence in their response. Eight word choices of equivalent letter count were displayed (0.40° × 0.95–3.14°), and participants indicated the correct word by clicking the computer mouse. On every trial, one of the response options was a related word. Next, three buttons ("very sure," "somewhat sure," and "guess") were presented, and participants clicked on the button that indicated their confidence in their preceding response. In the displays for both target identification and confidence rating, the options were arranged vertically, with equal spacing, and remained visible until a response was entered. Feedback regarding identification accuracy was then indicated by two colored rectangles (green = correct; red = incorrect) representing performance on the context (top) and target (bottom) words; if there were eye movements or blinks, this was noted in red text, and then the trial ended.

Electrophysiological Recording and Analysis

Continuous electroencephalogram (EEG) was collected (200-Hz sampling rate; 0.01- to 80-Hz band-pass filter; vertex reference) using a 128-channel Geodesic Sensor Net (Tucker, 1993) with individual electrode impedances kept below 50 to 75 kΩ; data were stored on a Power Mac G4 computer. EEG was processed off-line using Net Station 4.1.2 (Electrical Geodesics, Inc., Eugene, OR), beginning with a 40-Hz low-pass filter. Segmentation was time-locked to onset of the target display (segment epoch: −100 to 1,000 ms). All the segments within each experimental condition were averaged, rereferenced to the average reference, and baseline corrected for the 100-ms interval prior to target onset.

RESULTS AND DISCUSSION

Accuracy in target identification was analyzed using a 2 (mask offset: simultaneous vs. delayed) × 2 (word-pair relatedness: related vs. unrelated) repeated measures analysis of variance (ANOVA). The main effects of both offset, $F(1, 11) = 51.40$, $p_{rep} = .99$, $\eta_p^2 = .82$, and relatedness, $F(1, 11) = 31.21$, $p_{rep} = .99$, $\eta_p^2 = .74$, were significant, as was their interaction, $F(1, 11) = 15.32$, $p_{rep} = .99$, $\eta_p^2 = .58$. Separate repeated measures ANOVAs for the two relatedness conditions revealed significant OSM for both related word pairs (M = 17.88%), $F(1, 11) = 29.03$, $p_{rep} = .99$, $\eta_p^2 = .72$, and unrelated word pairs (M = 37.92%), $F(1, 11) = 44.40$, $p_{rep} = .99$, $\eta_p^2 = .80$, the latter exhibiting greater susceptibility to masking (see Fig. 2).

N400 amplitude was defined as the average amplitude for electrodes 5, 6, 7, 12, 81, 106, 107, and vertex (chosen on the basis of maximum peak amplitude in the simultaneous-offset condition) in a window from 355 to 755 ms. Figures 3a and 3b show N400 amplitude as a function of mask offset and word-pair relatedness. These data were analyzed in a 2 (mask offset) × 2 (relatedness) repeated measures ANOVA, which revealed a nonsignificant main effect of mask offset, $F(1, 11) = 2.54$, n.s.; a significant main effect of relatedness, $F(1, 11) = 18.43$, $p_{rep} = .99$, $\eta_p^2 = .63$; and, critically, a significant interaction, $F(1, 11) = 10.92$, $p_{rep} = .97$, $\eta_p^2 = .50$, indicating that the effect of relatedness on the N400 depended on mask offset. Separate repeated measures ANOVAs for the two mask-offset conditions revealed a relatedness effect (unrelated condition minus related condition) in the simultaneous-offset condition (M = −2.29 μV), $F(1, 11) = 23.72$, $p_{rep} = .99$, $\eta_p^2 = .68$, but not in the delayed-offset condition (M = −0.45 μV), $F(1, 11) = 1.45$, n.s. (see Fig. 3c).

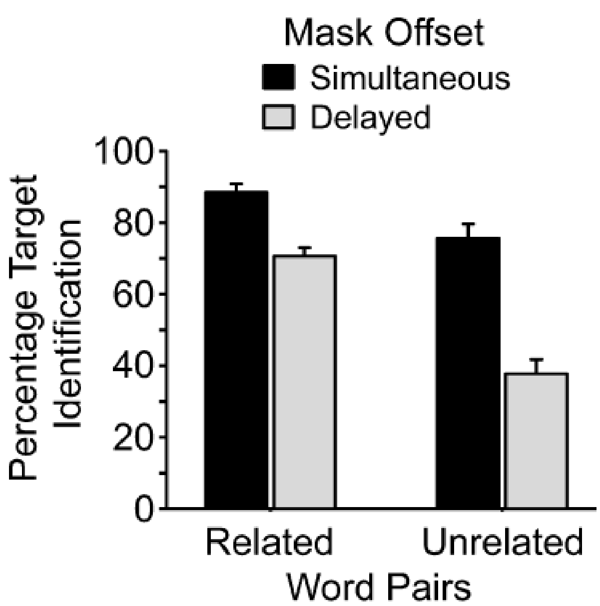

Fig. 2. Mean percentage of trials in which the target word was identified correctly as a function of mask offset (simultaneous vs. delayed) and word-pair relatedness (related vs. unrelated). Error bars represent 1 repeated measures *SEM* within each relatedness condition (Masson & Loftus, 2003).

The absence of a relatedness effect on the N400 when offset of the mask was delayed may appear to be at odds with the behavioral data showing a large relatedness effect in this condition (see Fig. 2). However, at least part of the behavioral relatedness effect in the delayed-offset condition appears to have been due to a change in guessing strategies. When participants were incorrect on unrelated trials, they chose the related word option on 31.32% of the trials in the simultaneous-offset condition and on 53.52% of the trials in the delayed-offset condition, $F(1, 11) = 11.52$, $p_{rep} = .97$, $\eta_p^2 = .51$. This increase in willingness to guess the related word would have increased accuracy when the context and target words were, in fact, related and would have decreased accuracy when they were unrelated, artificially inflating the difference between the related and unrelated conditions. Furthermore, participants may have occasionally guessed on the basis of incomplete target information (i.e., detection of one letter's features could eliminate several foils). Because accurate performance in this case would not rely on word reading, little semantic processing or N400 would be expected.[2]

The delayed offset of the masking dots had similar effects on behavior and N400 amplitude: Participants were severely impaired in identifying the masked words and showed no differential N400 activity for unrelated versus related words. These results suggest that OSM largely precluded semantic processing of the target. This finding stands in sharp contrast to results obtained in the AB paradigm, in which Vogel et al. (1998) found a striking dissociation between behavior and N400 amplitude: The AB dramatically impaired T2 awareness, but the N400 was unaffected. That dissociation suggested that interference in the AB paradigm occurs relatively late in processing, perhaps at the stage in which information is consolidated into working memory; interference at that stage would affect an observer's ability to report the target but would not preclude semantic analysis. By contrast, the current study offers the first evidence that the interference observed in OSM may occur at an earlier stage of processing, located somewhere between the derivation of basic visual features (Woodman & Luck, 2003) and semantic processing. This distinction is important because early researchers

[2]Follow-up ERP analyses conducted separately for correct and incorrect delayed-offset trials revealed that the relatedness effects in these two accuracy conditions were statistically indistinguishable ($F < 1$); in addition, relatedness did not have a significant effect on the N400 on either correct or incorrect delayed-offset trials, $Fs(1, 11) = 1.29$ and 2.54, respectively, n.s. These analyses provide further evidence that behavioral relatedness effects were primarily achieved without semantic processing of the target.

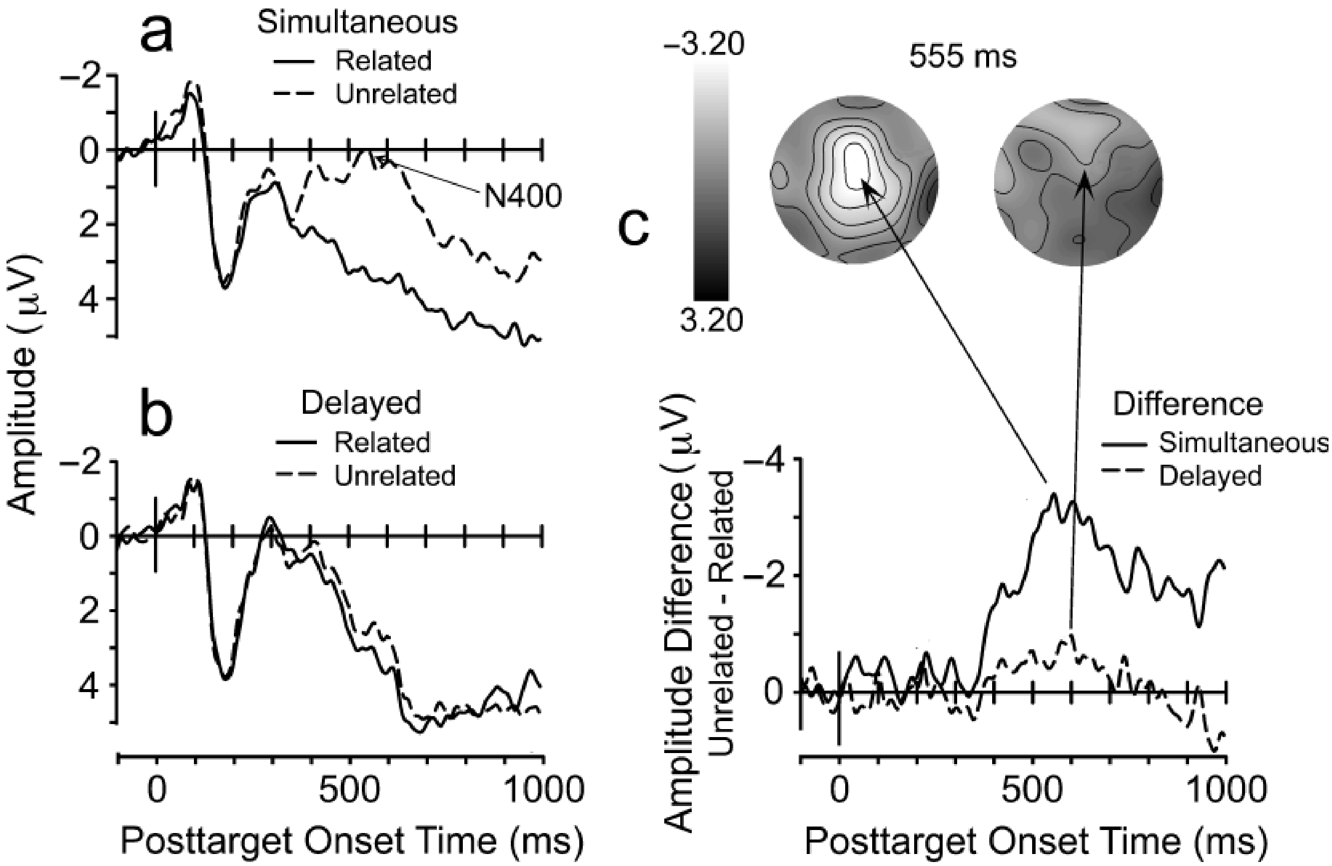

Fig. 3. Event-related potential (ERP) results. The graphs on the left present grand-averaged waveforms in response to presentation of related versus unrelated target words under simultaneous (a) and delayed (b) mask offset. ERP differences (unrelated minus related word pairs) for the two mask conditions are shown in (c), both as waveforms and as the corresponding distributions of the average N400 peak scalp amplitudes. Large full-color versions of these scalp distribution maps are available on the Web at http://hoffman.psych.udel.edu/OSM/.

(Brehaut, Enns, & Di Lollo, 1999; Giesbrecht & Di Lollo, 1998) proposed an essential role for substitution in the AB. However, Giesbrecht, Bischof, and Kingstone (2003) challenged this view, reporting that manipulating OSM strength failed to affect the AB. The current results generally support the latter claim that the AB and OSM are mediated by different mechanisms.

The present findings are consistent with the reentrance model proposed by Di Lollo et al. (2000), which holds that OSM is the result of a mismatch between the initial perception of the dots plus target as an integrated object and a later representation based on the persisting masking dots. The earlier representation is discarded in favor of the later one, resulting in awareness of the dots alone. This model emphasizes that interference is based on temporally extended cycles of bidirectional exchange of information between earlier and later visual processing areas and suggests that OSM reflects deficits in presemantic, visual representations. According to this model, target information requiring few cycles (e.g., simple features, as in Woodman & Luck, 2003) may escape masking, whereas information requiring additional cycles (i.e., semantics) may not (Enns & Di Lollo, 2000). This study provides a coarse estimate of the locus of impairment caused by OSM; further research using similar methods may provide more fine-grained information about the particular processing stages responsible for the OSM phenomenon.

Acknowledgments—This research was supported by National Science Foundation grants (MRI9977628 and BCS 0117475) to J.E.H. We would like to thank Susan Zup, Matthew Doran, Paul Quinn, James Cutting, Geoffrey Woodman, and an anonymous reviewer for their comments and suggestions on earlier drafts of this work.

REFERENCES

Alpern, M. (1953). Metacontrast. *Journal of the Optical Society of America*, *43*, 648–657.

Brehaut, J.C., Enns, J.T., & Di Lollo, V. (1999). Visual masking plays two roles in the attentional blink. *Perception & Psychophysics*, *61*, 1436–1448.

Di Lollo, V., Bischof, W.F., & Dixon, P. (1993). Stimulus-onset asynchrony is not necessary for motion perception or metacontrast masking. *Psychological Science*, *4*, 260–263.

Di Lollo, V., Enns, J.T., & Rensink, R.A. (2000). Competition for consciousness among visual events: The psychophysics of reentrant visual processes. *Journal of Experimental Psychology: General*, *129*, 481–507.

Enns, J.T., & Di Lollo, V. (1997). Object substitution: A new form of masking in unattended visual locations. *Psychological Science*, *8*, 135–139.

Enns, J.T., & Di Lollo, V. (2000). What's new in visual masking? *Trends in Cognitive Sciences*, *4*, 345–352.

Giesbrecht, B., Bischof, W.F., & Kingstone, A. (2003). Visual masking during the attentional blink: Tests of the object substitution hypothesis. *Journal of Experimental Psychology: Human Perception and Performance*, *29*, 238–258.

Giesbrecht, B., & Di Lollo, V. (1998). Beyond the attentional blink: Visual masking by object substitution. *Journal of Experimental Psychology: Human Perception and Performance*, *24*, 1454–1466.

Jiang, Y.H., & Chun, M.M. (2001). The spatial gradient of visual masking by object substitution. *Vision Research*, *41*, 3121–3131.

Kutas, M., & Hillyard, S.A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science*, *207*, 203–205.

Lamme, V.A.F., & Roelfsema, P.R. (2000). The distinct modes of vision offered by feedforward and recurrent processing. *Trends in Neurosciences*, *23*, 571–579.

Lamme, V.A.F., Zipser, K., & Spekreijse, H. (2002). Masking interrupts figure-ground signals in V1. *Journal of Cognitive Neuroscience*, *14*, 1044–1053.

Lleras, A., & Moore, C.M. (2003). When the target becomes the mask: Using apparent motion to isolate the object-level component of object substitution masking. *Journal of Experimental Psychology: Human Perception and Performance*, *29*, 106–120.

Luck, S.J., & Hillyard, S.A. (1994). Spatial filtering during visual search: Evidence from human electrophysiology. *Journal of Experimental Psychology: Human Perception and Performance*, *20*, 1000–1014.

Luck, S.J., Vogel, E.K., & Shapiro, K.L. (1996). Word meanings can be accessed but not reported during the attentional blink. *Nature*, *383*, 616–618.

Masson, M.E.J., & Loftus, G.R. (2003). Using confidence intervals for graphically based data interpretation. *Canadian Journal of Experimental Psychology*, *57*, 203–220.

Moore, C.M., & Lleras, A. (2005). On the role of object representations in substitution masking. *Journal of Experimental Psychology: Human Perception and Performance*, *31*, 1171–1180.

Neill, W.T., Hutchison, K.A., & Graves, D.F. (2002). Masking by object substitution: Dissociation of masking and cuing effects. *Journal of Experimental Psychology: Human Perception and Performance*, *28*, 682–694.

Raymond, J.E., Shapiro, K.L., & Arnell, K.M. (1992). Temporary suppression of visual processing in an RSVP task: An attentional blink? *Journal of Experimental Psychology: Human Perception and Performance*, *18*, 849–860.

Rolke, B., Heil, M., Streb, J., & Hennighausen, E. (2001). Missed prime words within the attentional blink evoke an N400 semantic priming effect. *Psychophysiology*, *38*, 165–174.

Shapiro, K., Driver, J., Ward, R., & Sorensen, R.E. (1997). Priming from the attentional blink: A failure to extract visual tokens but not visual types. *Psychological Science*, *8*, 95–100.

Tucker, D.M. (1993). Spatial sampling of head electrical fields: The Geodesic Sensor Net. *Electroencephalography and Clinical Neurophysiology*, *87*, 154–163.

Turvey, M.T. (1973). Peripheral and central processes in vision: Inferences from an information-processing analysis of masking with patterned stimuli. *Psychological Review*, *80*, 1–52.

Vogel, E.K., Luck, S.J., & Shapiro, K.L. (1998). Electrophysiological evidence for a postperceptual locus of suppression during the attentional blink. *Journal of Experimental Psychology: Human Perception and Performance*, *24*, 1656–1674.

Woodman, G.F., & Luck, S.J. (2003). Dissociations among attention, perception and awareness during object-substitution masking. *Psychological Science*, *14*, 605–611.

(RECEIVED 2/20/06; REVISION ACCEPTED 5/1/06; FINAL MATERIALS RECEIVED 5/10/06)