Does Consistent Scene Context Facilitate Object Perception?

Andrew Hollingworth and John M. Henderson
Michigan State University

The conclusion that scene knowledge interacts with object perception depends on evidence
that object detection is facilitated by consistent scene context. Experiment 1 replicated the
I. Biederman, R. J. Mezzanotte, and J. C. Rabinowitz (1982) object-detection paradigm.
Detection performance was higher for semantically consistent versus inconsistent objects.
However, when the paradigm was modified to control for response bias (Experiments 2 and 3)
or when response bias was eliminated by means of a forced-choice procedure (Experiment 4),
no such advantage obtained. When an additional source of biasing information was eliminated
by presenting the object label after the scene (Experiments 3 and 4), there was either no effect
of consistency (Experiment 4) or an inconsistent object advantage (Experiment 3). These
results suggest that object perception is not facilitated by consistent scene context.

To what degree is perception affected by our knowledge of the world? This question has historically been central in theories of perception and cognition. For example, the apparent role of semantic constraint in visual recognition led to the emergence of so-called New Look psychology (Bruner, 1957, 1973). In cognitive psychology, the effects of contextual expectations on perception have been couched in terms of debates between bottom-up versus top-down pattern recognition (Neisser, 1967) and modular versus interactive perception (e.g., Fodor, 1983; Pylyshyn, 1980; Rumelhart, McClelland, & the PDP Research Group, 1986). More recently, the effect of higher-level knowledge on perception has become critical in discussions of the role of re-entrant neural pathways in the early cortical processing of visual stimulation (Barlow, 1994; Churchland, Ramachandran, & Sejnowski, 1994; Kosslyn, 1994; Mumford, 1994). In the present study, we explored the following version of the context question: How is the identification of a visual object affected by the meaning of the real-world scene in which that object appears? This question is important because it

directly addresses the influence of our knowledge and beliefs about meaningful relationships in the world on our perception of the visual environment.

For the purposes of this study, we define *object identification* from the perspective of current computational theories (Biederman, 1987; Bülthoff, Edelman, & Tarr, 1995; Marr, 1982; Marr & Nishihara, 1978; Ullman, 1996). At a general level of description, these theories assume two processing stages in object identification. First, the retinal image is transformed into a perceptual description that is compatible with a set of memory descriptions. This first stage can be further broken down into two sub-stages, an early stage of visual analysis that translates the current pattern of retinal stimulation into perceptual primitives, and an additional stage that uses these primitives to produce descriptions of the object tokens in the scene. Second, object descriptions are matched to stored long-term memory descriptions of object types, leading to entry-level recognition. When a match is found, identification has occurred, and information stored in memory about that object type, such as its identity, whether it is good to eat, and so on becomes available. In the following experiments, we will consider the activation of an entry-level label for an object stimulus as evidence of the successful completion of the matching stage of object identification.

The hypothesis we set out to test was that the identification of a real-world object is facilitated when that object is semantically consistent rather than inconsistent with the scene in which it appears (Biederman, 1981; Biederman, Mezzanotte, & Rabinowitz, 1982; Boyce & Pollatsek, 1992; Boyce, Pollatsek, & Rayner, 1989; Friedman, 1979; Kosslyn, 1994; Metzger & Antes, 1983; Palmer, 1975a; Ullman, 1996; see Henderson & Hollingworth, in press, for a review). The strongest evidence supporting the view that consistent scene context facilitates object identification comes from the object-detection paradigm introduced by Biederman and colleagues (Biederman, 1981; Biederman et al., 1982). In this paradigm, participants were asked to determine whether a pre-specified object appeared within a briefly presented scene at a particular location. During each

trial, a label naming an object was presented until the participant was ready to continue, followed by a line-drawing of a natural scene presented for 150 ms, followed by a pattern mask with an embedded location cue. The pattern mask remained on the screen until the participant pressed one of two buttons to indicate whether the object named by the label had or had not appeared in the scene at the cued location.[1] In target-present trials, the target label named the cued object. In catch trials, the target label named an object that did not appear in the scene. We will refer to this paradigm as the *original object-detection paradigm*.

The primary contextual manipulation in the original object-detection paradigm was the relationship between the object presented at the cued location (the cued object) and the scene in which that object appeared. *Base* scenes contained a cued object that was consistent with the scene, and the scene contained no other objects that violated scene context. *Violation* scenes contained a cued object that was inconsistent with the scene along one or more dimensions, including episodic probability, position, size, support, and interposition (whether the object occluded objects behind it or was transparent). For the purposes of this article, we will focus on cases of probability violation (i.e., the semantic consistency between object and scene). The most conservative hypothesis regarding the information contained in a scene concept is that it specifies the object types typically found in the scene (Mandler & Johnson, 1976; Mandler & Parker, 1976). Therefore, manipulations of object probability provide the most direct means to investigate the influence of scene meaning on object perception.

Biederman et al. (1982) found that detection performance was best when the cued object did not violate any of the constraints imposed by scene meaning. They reported poorer performance across all violation dimensions, with compound violations (e.g., probability and support) producing even greater decrements. Importantly, violations of semantic relationships were found to be as disruptive as violations of structural relationships (e.g., support or interposition), suggesting that semantic relationships can be accessed very rapidly and can then interact with the initial perceptual analysis of an object. The poorer performance in violation conditions held across percentage correct performance, sensitivity (d'), and reaction time. These measures, however, are not equally valid for investigating the influence of scene context on object perception. Reaction time in that study cannot be taken as strong evidence for the facilitated detection of consistent objects, as there were a significant number of errors, causing more than 30% of the trials to be excluded from the analysis. In general, it is not clear that reaction time is a good measure of object perception processes, as it may be influenced by post-identification factors, such as response generation (Henderson, 1992).[2] In addition, percentage correct performance in target-present trials does not provide a reliable measure of object-detection performance, as participants demonstrated significant response biases that varied with the consistency of the target label with the scene. Thus, the conclusion that consistent scene context facilitates object perception rests most firmly

on the finding that detection sensitivity (d') was higher in base conditions than in violation conditions. This sensitivity result has been replicated by Boyce et al. (1989) using a similar object-detection paradigm.

The results of the Biederman et al. (1982) study have led to two general conclusions about the perception of objects in natural scenes. First, scene meaning, including information about the semantic relationship between scene and object types, can be accessed very quickly. Such early activation is a necessary condition for context effects, as contextual constraints must be active early enough to influence the perception of object stimuli. This conclusion is supported by a number of studies demonstrating that the information necessary to identify natural scenes can be obtained in less than 150 ms (Antes, Penland, & Metzger, 1981; Biederman, 1972; Biederman, Glass, & Stacy, 1973; Loftus, Nelson, & Kallman, 1983; Potter, 1976; Schyns & Oliva, 1994). Second, stored knowledge about scenes and the objects likely to appear in them can be used to facilitate the construction of perceptual descriptions of consistent objects. Taken together, these two hypotheses combine to form the *perceptual schema model* of scene context effects (Biederman, 1981; Biederman et al., 1982; Palmer, 1975b; see Henderson, 1992, for further discussion). The perceptual schema model proposes that the stored representation of a scene type contains information about the objects that form that type. The early activation of this information can be used to facilitate the perceptual analysis (e.g., the encoding of features or generation of a perceptual description) of objects that are consistent with the semantic constraints imposed by the scene.

In addition to the perceptual schema model, two other models of scene and object processing can account for the facilitated detection of consistent objects in scenes. First, a *priming model* of scene context effects (Friedman, 1979; Kosslyn, 1994; Palmer, 1975a) places the locus of contextual influence at the matching stage of object identification, when the perceptual description of an object token is compared to long-term memory representations of object types. According to the priming model, the recognition of a scene serves to prime the stored representations of object types consistent with that scene (i.e., the activation levels of stored, consistent object representations are raised closer to a threshold value). As a result, relatively less perceptual information needs to be encoded to bring a stored, consistent

object representation to the threshold value indicating that a match has been found.[3]

Second, facilitated detection of consistent objects in scenes can be accounted for by an *interactive activation model* similar to that proposed by McClelland and Rumelhart (1981) for word and letter recognition (Boyce & Pollatsek, 1992; Boyce et al., 1989; Metzger & Antes, 1983). In this model, scenes correspond to the word level in the network and objects to the letter level. These two levels mutually constrain each other, facilitating the perception of objects consistent with scene meaning and inhibiting the perception of inconsistent objects. In addition, partial activation at the object level could act to constrain the encoding of perceptual features consistent with that object type, though no interactive activation model of scene context effects has, as yet, specifically included that level of interaction.

Each of these models of the influence of scene knowledge on object perception predicts that perception of an object should be facilitated when that object is consistent with the scene in which it appears. We will therefore refer to them as *contextual facilitation models* of object perception in scenes.

Concerns With the Object-Detection Paradigm

Support for contextual facilitation models rests almost entirely on results from object-detection experiments that show detection benefits for consistent objects versus inconsistent objects under brief presentation conditions (Biederman et al., 1982; Boyce et al., 1989). However, a number of general concerns have been raised regarding the original object-detection paradigm (De Graef, Christiaens, & d'Ydewalle, 1990; De Graef & d'Ydewalle, 1995; Henderson, 1992). These concerns revolve around two central issues. First, the original object-detection paradigm may not have adequately controlled participant response bias. Second, the presentation of the target label prior to the scene and the location cue following the scene may have provided additional sources of information that influenced detection performance.

Catch Trial Design and the Calculation of Sensitivity

In the original object-detection paradigm (Biederman et al., 1982), detection sensitivity (d') was calculated using the percentage correct rate in target-present trials (the hit rate) and the error rate in catch trials (the false-alarm rate). Table 1 summarizes the design of the target-present and catch trials for Biederman et al. (1982, Experiment 1). One concern with this design is that the method of calculating detection sensitivity did not adequately control for participants' bias to respond "yes" more often when the catch trial label was semantically consistent versus semantically inconsistent with the subsequent scene. If participants attempt to detect an absent horse in a farmyard, for example, they should be biased to respond "yes," as a horse is generally likely to appear there. If participants attempt to detect an absent television in the farmyard, however, they should show less bias to respond "yes," as there is contextual information arguing against a television's presence. This pattern of bias

Table 1

Summary of the Target-Present and Catch Trial Design in Biederman et al. (1982, Experiment 1) and in Experiments 1–3 for a Sample Trial Presenting a Farmyard Scene

Trial	Semantic consistency manipulation	
	Consistent (base)	Inconsistent (probability violation)
Biederman et al. (1982)		
Target-present		
Cued object	chicken	mixer
Target label	"chicken"	"mixer"
Catch		
Cued object	pig	mixer
Target label	70% consistent ("horse"), 30% inconsistent ("television")	70% consistent ("horse"), 30% inconsistent ("television")
Experiment 1		
Target-present		
Cued object	chicken	mixer
Target label	"chicken"	"mixer"
Catch		
Cued object	chicken	mixer
Target label	50% consistent ("horse"), 50% inconsistent ("television")	50% consistent ("horse"), 50% inconsistent ("television")
Experiments 2 and 3		
Target-present		
Cued object	chicken	mixer
Target label	"chicken"	"mixer"
Catch		
Cued object	mixer	chicken
Target label	"chicken"	"mixer"

should lead to higher false-alarm rates in consistent target label catch trials and lower false-alarm rates in inconsistent target label catch trials. Biederman et al. found precisely this effect: False-alarm rates were higher when the target label

---

[3] Friedman (1979) recorded eye movements while participants viewed scenes in preparation for a difficult memory test. The *duration of the first fixation* (i.e., the total duration the eyes were fixated on the object the first time it was entered, now referred to as *first pass gaze duration*) was shorter for consistent versus inconsistent objects, a result that Friedman (1979) interpreted as support for a priming model of scene context effects. The difference in fixation duration was more than 300 ms, however, which is unlikely to be explained by perceptual factors alone (Biederman et al., 1982; Henderson, 1992). For example, participants may have looked longer at a semantically inconsistent object to integrate the already identified object into a conceptual representation in which it was incongruous. In addition, the instructions to prepare for a difficult memory test may have caused participants to consciously create an association between the inconsistent object and the scene, leading to longer fixation durations (Biederman et al., 1982; Henderson, 1992).

was semantically consistent versus inconsistent with the subsequent scene.

To eliminate this sort of bias from measures of object-detection performance, detection sensitivity in the base condition should be calculated using the hit rate in trials when the target label was consistent with the scene and the false-alarm rate in trials when the target label was also consistent with the scene. Similarly, detection sensitivity for target objects in the probability violation condition should be calculated using the hit rate in trials when the target label was inconsistent with the scene and the false-alarm rate in trials when the target label was also inconsistent with the scene. The Biederman et al. (1982) study, however, did not calculate sensitivity in this manner: The false-alarm rates in both base and violation conditions averaged across catch trials on which the target label was consistent and inconsistent with the subsequent scene (see Table 1). Thus, differences in response bias as a function of target label consistency were not necessarily controlled in the d' measure.

It is important to note that this method of computing d' may have overestimated the sensitivity rate in base conditions and underestimated the sensitivity rate in violation conditions. As discussed previously, the false-alarm rate was higher when the catch trial label was consistent versus inconsistent with the scene. By averaging across these two catch trial conditions for the purpose of calculating d', sensitivity in the base condition may have been artificially raised because the averaged false-alarm rate was lower than the false-alarm rate for consistent target label catch trials alone. Similarly, sensitivity in the probability violation condition may have been artificially lowered because the averaged false-alarm rate was higher than the false-alarm rate for inconsistent target label catch trials alone. Thus, the Biederman et al. (1982) method of calculating d' very likely produced an exaggerated sensitivity advantage for the detection of consistent objects.[4]

A second concern with the design of the original object-detection paradigm is that participants did not attempt to detect the same object in the catch trials as in the corresponding target-present trials. Thus, detection sensitivity for a particular object was based on the correct detection of that object in the scene and the false detection of an entirely different object that was not in the scene. Signal detection theory, however, requires that sensitivity measures be calculated using the correct detection of a particular signal when it is present and the false detection of the *same* signal when it is not present (Green & Swets, 1966). For example, suppose that the consistent cued object appearing in a kitchen scene were a stove, and the consistent target label in the catch trial were "bread box." The hit rate would reflect the correct detection of a stove. The false-alarm rate, however, would reflect the false detection of a bread box. Because a bread box is less likely to appear in a kitchen scene than a stove, the false-alarm rate would be artificially low, and the resulting sensitivity estimate would be artificially high. This example demonstrates the importance of requiring participants to detect the same object on corresponding target-present and catch trials. The object-detection experiments conducted to date (Biederman et al., 1982; Boyce et al.,

1989) have not met this criterion (see De Graef & d'Ydewalle, 1995), and thus the sensitivity measures in those experiments must be interpreted with caution.

Target Label Preview

The second concern with the original object-detection paradigm involves the presentation of the target label before scene viewing. There are two potential problems with such a design. First, participants may have used the identity of the target object to guide their search in the subsequent scene (Henderson, 1992). This strategy could have lead to a consistent object-detection advantage if the spatial positions of consistent objects were more predictable than the positions of inconsistent objects. For example, participants may have known where to find a horse in a farmyard but would not necessarily have known where to find a television in a farmyard. Supporting this intuition, we have recently demonstrated that viewers can more quickly locate semantically consistent versus inconsistent objects in a free-viewing, visual search task (Henderson, Weeks, & Hollingworth, in press). The information provided by the target label preview may have been particularly helpful in the Boyce et al. (1989) experiments. The scenes used in that study contained only five discrete objects presented against a very simple background. Thus, participants had to search through only a small number of objects during scene presentation. If the positions of consistent objects were more predictable than the positions of inconsistent objects, an advantage for the detection of consistent objects would be expected.

Second, the preview of the target label (e.g., "mixer") may have led participants to expect the subsequent presentation of a certain scene type (in this case, a kitchen). The generation of such expectations would have been particularly likely in the Biederman et al. (1982) study, because 70% of the target labels were consistent with the subsequent scene. When the target label specified an object inconsistent with the subsequent scene, however, the discontinuity between the expected and presented scene may have interfered with perceptual processing, to the detriment of inconsistent object detection.

Location Cue

The final concern regarding the original object-detection paradigm is that the location cue may have provided information useful to post-perceptual guessing. The position of the cue relative to the scene could have provided evidence concerning the types of objects likely to be found at that position (Henderson, 1992). A cue marking a position high in a living room scene, for example, would constrain the objects that could have appeared there to clocks, pictures, curtains, etc. Such information would be useful when attempting to detect a consistent object but not as useful

---

[4] It is important to note that the Boyce et al. (1989) replication of this paradigm is not subject to this criticism, as the consistency of the catch trial labels was controlled. However, the Boyce et al. study is subject to the remaining criticisms.

when attempting to detect an inconsistent object, because the spatial position of an inconsistent object is less predictable.

The Present Study

Results from object-detection experiments are the primary support for the widely held view that consistent scene context facilitates object identification. It is therefore important to assess the previous concerns experimentally. To do this, we conducted four experiments. Experiment 1 attempted to replicate the original Biederman et al. (1982) paradigm. Experiment 2 tested whether the consistent object advantage in the original object-detection paradigm was due to the inadequate control of response bias. The original paradigm was modified so that participants attempted to detect the same object on corresponding target-present and catch trials. Experiment 3 investigated the influence of presenting the target label before the scene by manipulating whether the label appeared before or after the scene. In addition, the location cue was eliminated from the third experiment to investigate whether its presence may have affected performance. Experiment 4 employed a forced-choice procedure to investigate the influence of scene context on object perception independently of response bias. In all four experiments, contextual facilitation models predict better detection of objects that are consistent with a scene versus objects that are inconsistent, because they propose that stored scene knowledge of the types of objects likely to be found in a scene facilitates the identification of those objects.

Experiment 1

The purpose of Experiment 1 was to replicate the original Biederman et al. (1982) paradigm. We felt that replication was necessary as a baseline against which to compare the results from subsequent experiments in which we modified the original paradigm. Figure 1 illustrates the main aspects of the paradigm. The basic design was the same as that of Biederman et al. (1982, Experiment 1).[5] A target label was presented for 1500 ms, followed by a line drawing of a real-world scene for 200 ms, followed by a pattern mask with an embedded location cue. The participant's task was to determine whether the object named by the target label had or had not appeared in the scene at the cued location.

The key manipulation in Experiment 1 was the semantic consistency between the cued object and the scene in which it appeared. Semantically consistent cued objects were likely to appear in the scene (e.g., a chicken in a farmyard); semantically inconsistent cued objects were unlikely to appear in the scene (e.g., a mixer in a farmyard). In half of the trials, the target label named a semantically consistent object, and in the other half the target label named a semantically inconsistent object. Half of the trials presented a scene that contained the target object (target-present trials), and half of the trials presented a scene that did not (catch trials). In the target-present trials, the target label named the
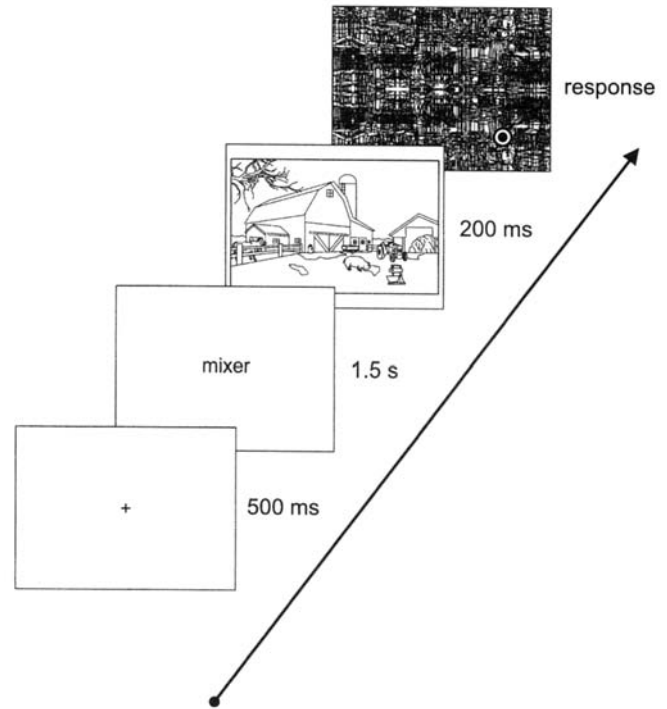

Figure 1. Schematic illustration of a trial in Experiment 1.

cued object. In the catch trials, the target label named an object that did not appear in the scene. As in the original object-detection paradigm, there was no relationship between the semantic consistency of the target label on

---

[5] We made a number of modifications to the original Biederman et al. (1982, Experiment 1) paradigm. First, we limited the cued object consistency manipulations to base (semantically consistent) and probability violation (semantically inconsistent). Second, we did not include a bystander condition. Third, we used a paired-scene design to control for scene-specific factors such as cued object distance from fixation and lateral masking. Fourth, we presented the target label for 1500 ms rather than for a participant-determined duration. Fifth, we used a scene presentation duration of 200 ms rather than 150 ms. The 200 ms presentation duration prevents eye movements during scene presentation, limiting viewing to a single glance of the scene, but reduces the possibility of floor effects. Sixth, in the Biederman et al. experiment, 30% of the catch trial target labels were inconsistent with the subsequent scene, because 30% of the labels in the target-present trials were inconsistent. In Experiment 1, 50% of the catch trial target labels were inconsistent with the subsequent scene, because 50% of the labels in the target-present trials were inconsistent. Finally, the probability violation catch trials in the Biederman et al. paradigm did not have the same design as the catch trials in the base condition. For each scene in the probability violation condition, the same object was cued in the catch trials as in the target-present trials. For each scene in the base condition, however, a different object was cued in the catch trials than in the target-present trials (see Table 1). Thus, we chose to make the two consistency conditions in Experiment 1 equivalent by always cueing the same object in target-present and catch trials for each scene in each consistency condition. This cueing manipulation is the same as that employed by Boyce et al. (1989).

corresponding target-present and catch trials: For each cued object consistency condition, half of the catch trials presented a consistent target label and half an inconsistent target label. Note that, as in the original object-detection paradigm, the catch trial semantic consistency manipulation is defined by the cued object appearing in the scene and not by the consistency of the object the participant is attempting to detect. Table 1 summarizes the design of the target-present and catch trials for Experiment 1.

Method

Participants. Twenty-four Michigan State University undergraduate students participated in the experiment for course credit. All participants had normal or corrected-to-normal vision. The participants were naive with respect to the hypotheses under investigation.

Stimuli. Twenty scenes and 20 cued objects were used as stimuli. The stimuli were generated from photographs of natural scenes. Fourteen scenes were generated from those used by van Diepen and De Graef (1994), and the other 6 scenes were generated from photographs taken in the East Lansing, Michigan, area. In both cases, the main contours of the scenes were traced using commercial software to create gray-scale line drawings. The images generated from the two sources were not distinguishable. Semantically consistent cued objects for each scene were also created by digitally tracing scanned images. These objects were created separately from the scenes. The 20 scenes were paired, and the semantically inconsistent conditions were created by swapping objects across scenes. For example, a mixer was the semantically consistent cued object in a kitchen scene, and a live chicken was the consistent cued object in a farmyard scene. These objects were swapped across scenes so that the mixer was the semantically inconsistent object in the farmyard scene, and the chicken was the inconsistent object in the kitchen scene. Figure 2 shows an example of a stimulus scene and the semantic consistency manipulation.

Because each scene was to be presented a total of eight times during the experiment, two cued object positions were chosen within each scene to minimize participants' ability to predict the object's location. Each position was chosen as a place within the scene where the consistent cued object might reasonably appear. Both the consistent and inconsistent cued objects appeared in the same two positions. In a number of scenes, the two object positions required using two different sizes of each object (e.g., a fire hydrant placed in two positions on a receding sidewalk). In such a case, the paired scene also employed two different sizes of each object. The percentage change in the size of each object was equated and was the same in both of the paired scenes. As a consequence of this paired-scene design, each scene served as a control for its partner, reducing the influence of such factors as object size, eccentricity, and lateral masking.

All scene and object manipulations were conducted using commercially available software. The scenes subtended a visual angle of 23 degrees (width) by 15 degrees (height) at a viewing distance of 64 cm. Cued objects subtended about 2.75 degrees on average (range = 1.25 to 4.92 degrees). All images were displayed as gray-scale contours on a white background at a resolution of 800 × 600 pixels × 16 levels of gray. Gray-scale was used for anti-aliasing so that the contours appeared smooth and sharp.

Target labels were created using lower-case, 24-point, anti-aliased Arial font. Labels for target-present trials named the cued object. For catch trials, one consistent and one inconsistent label were chosen for each scene. Each of these labels named an object that never appeared in the scene.
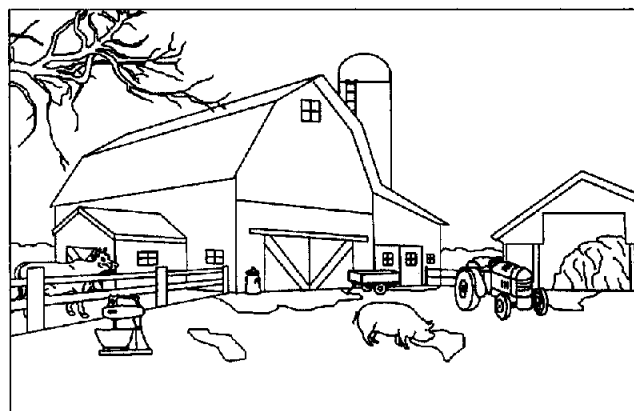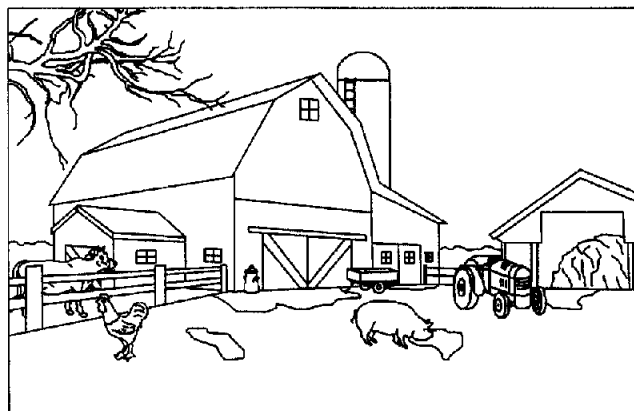

Figure 2. An example of the type of scene used and the cued object semantic consistency manipulation. The top scene contains a semantically consistent cued object (chicken), and the bottom scene contains a semantically inconsistent cued object (mixer). This farmyard scene was paired with a kitchen scene in which the mixer was consistent and the chicken inconsistent.

The pattern mask presented after the scene consisted of overlapping line segments, curves, and angles, and was slightly larger than the scene stimuli. The scenes were completely obliterated when presented simultaneously with the pattern mask. The location cue appearing within the pattern mask was a thick circle containing a dot, subtending 1.4 degrees.

Apparatus. The stimuli were displayed on a flat-screen SVGA monitor with a 100-Hz refresh rate. Responses were collected with a button box connected to a dedicated input–output (I–O) board. Depression of a button stopped a millisecond clock on the I–O board. The display and I–O systems were interfaced with a 486-based microcomputer that controlled the experiment.

Procedure. Participants were tested individually. The experimenter first explained that the task would be to determine whether the object named by a label was present at a marked location in a briefly displayed scene. The participant was then seated in front of a computer monitor, with one hand resting on a button labeled "yes" and the other on a button labeled "no." Viewing distance was maintained by a forehead rest.

During each experimental trial, participants saw a fixation cross and a prompt instructing them to press a pacing button to begin the trial. Once the participant pressed the button, the fixation cross remained on the screen for an additional 500 ms, followed by a

blank screen containing a target label for 1500 ms, followed by the presentation of the scene for 200 ms, followed by a pattern mask containing an embedded location cue. There was no delay (i.e., the inter-stimulus interval was zero) between each display. The pattern mask remained in view until the participant pressed the left (yes) button to indicate that the object named by the target label had appeared in the scene at the cued location or the right (no) button to indicate that the target object had not appeared at the cued location. After the response, there was a 4-s delay while the stimuli for the next trial were loaded into video memory, and then the prompt for the next trial appeared.

Participants took part in a practice block of 16 trials (2 cued object consistency conditions × 2 target object presence conditions × 2 cued object positions × 2 scenes). The two scenes used in the practice block were not used in the experimental trials. After the practice trials, the experimenter answered any questions the participant had about the procedure, and the participant proceeded to the experimental trials.

Each participant saw 160 experimental trials that were produced by a within-participant factorial combination of 2 cued object consistency conditions × 2 target object presence conditions × 2 cued object positions × 20 scenes. The position of the cued object and the consistency of the catch trial label were counterbalanced between a pair of scenes and completely counterbalanced as a between-participants factor. Because the cued object position factor was not of theoretical interest, the two levels of that factor were combined in the statistical analyses. Each participant saw all 160 trials in a different random order. The entire session lasted approximately 45 min.

Results

In the following analyses, two measures were used to assess object-detection performance. First, we report percentage correct hits for the target-present trials and percentage correct rejections for the catch trials. Second, because reliable response biases were present, we report A', a nonparametric measure of sensitivity (Grier, 1971). A' can be interpreted as equivalent to percentage correct in a forced-choice procedure. A' was computed using the rate of correct responses in target-present trials (the hit rate) and the rate of errors in catch trials (the false-alarm rate). For the purpose of replicating the Biederman et al. (1982) paradigm, we did not separate the catch trial data as a function of the semantic consistency of the target label. Thus, the A' results reported in this section are based on the mean false-alarm rate in each cued object consistency condition, averaging across catch trials in which the target label was consistent and inconsistent with the subsequent scene.

Percentage correct analysis. Mean percentage correct as a function of cued object consistency and target object presence is presented in Table 2. First, there was a reliable main effect of target object presence, $F(1, 23) = 12.79$, $MSE = .0506$, $p < .005$. Participants responded correctly 67.9% of the time when the target object was present and 79.5% of the time when the target object was absent. There was also a main effect of cued object consistency, $F(1, 23) = 31.57$, $MSE = .0141$, $p < .001$, with better performance when the cued object was consistent (78.5%) than when it was inconsistent (68.9%) with the scene. Finally, there was a reliable interaction between cued object consistency and

target object presence, $F(1, 23) = 19.42$, $MSE = .0149$, $p < .001$. Simple effects tests indicated that the hit rate for scenes containing a consistent cued object (76.6%) was higher than that for scenes containing an inconsistent cued object (59.2%), $F(1, 23) = 31.71$, $MSE = .0229$, $p < .001$. This pattern was not present in the catch trials; the correct rejection rate was the same when the cued object was consistent (80.4%) versus inconsistent (78.5%), $F(1, 23) = 1.38$, $MSE = .0061$, $p > .25$. To summarize, although the hit rate was higher for consistent versus inconsistent cued objects, the false-alarm rates were not different, 19.6% and 21.5% respectively.

A' analysis. Mean A' for each cued object consistency condition is shown in Table 2. Detection sensitivity was reliably higher in scenes containing a consistent cued object (.861) than in scenes containing an inconsistent cued object (.775), $F(1, 23) = 20.50$, $MSE = .0043$, $p < .001$.

Discussion

Experiment 1 replicated the principal result of the original object-detection paradigm. We found that object-detection performance was higher for scenes containing a consistent cued object than for scenes containing an inconsistent cued object. This experiment demonstrates that our scenes and semantic consistency manipulation are sufficient to replicate the basic consistent object contextual facilitation effect. The results of this experiment provide a baseline that can be used to assess the results of Experiments 2–4, in which the original paradigm will be modified to address the concerns specified in the Introduction.

In Biederman et al. (1982), false-alarm rates were higher when the catch trial label was consistent versus inconsistent with the scene. To investigate whether such biases were present in the current experiment, we conducted a full analysis of the Experiment 1 catch trial data. The two within-participants variables of cued object semantic consistency and target label semantic consistency were entered into an ANOVA. There was a reliable effect of target label semantic consistency, $F(1, 23) = 4.97$, $MSE = .0303$, $p < .05$, with a higher correct rejection rate for inconsistent target labels (83.4%) than consistent target labels (75.5%). As found in the Biederman et al. study, participants were more likely to falsely respond that an absent consistent target object was present than to respond that an absent inconsistent target object was present. This response bias was likely caused by participants adopting a higher standard of evidence to accept that an inconsistent object was present in the scene than that a consistent object was present.

Table 2
Mean Percentage Hits, Percentage Correct Rejections (Percentage False Alarms), and A' for Experiment 1

Cued object consistency	% hits	% correct rejections (% false alarms)	A'
Consistent	76.6	80.4 (19.6)	.861
Inconsistent	59.2	78.5 (21.5)	.775

As reported previously, the effect of cued object semantic consistency on catch trial performance did not approach reliability, $F(1, 23) = 1.38, MSE = .0061, p > .25$, nor was the interaction between cued object consistency and target label consistency reliable, $F < 1$. The absence of a cued object consistency effect on catch trial performance is intriguing. According to contextual facilitation models, consistent cued objects should be easier to identify than inconsistent cued objects. Thus, contextual facilitation models predict lower false-alarm rates when the cued object is consistent with the scene, because participants will be better able to determine that the cued object does not match the target label. No such effect of cued object consistency was found. It is important to note that a similar lack of an effect of cued object consistency on catch trial performance was reported by Biederman et al. (1982, Experiment 1). In that experiment, false-alarm rates were 16% in the base condition and 15% in the probability violation condition.

In summary, analysis of the catch trial data revealed that the semantic consistency of the target label reliably influenced performance, but the semantic consistency of the cued object had little influence on performance. Because the original object-detection paradigm (and our replication of that paradigm in Experiment 1) averaged across consistent and inconsistent target label catch trials when calculating false-alarm rates, it is likely that these experiments underestimated the false-alarm rate for base conditions and overestimated the false-alarm rate for violation conditions. This would result in artificially high sensitivity estimates in base conditions and artificially low sensitivity estimates in violation conditions. Overall, these data suggest that the consistent object facilitation effect observed in prior object-detection experiments may have been due, at least in part, to the fact that sensitivity measures did not control for response biases induced by the consistency of the target label.

Experiment 2

In Experiment 2 we modified the original object-detection paradigm to address our concerns about catch trial design and the method of calculating sensitivity. In this experiment, the target label in a catch trial named the same object as in a corresponding target-present trial (see Table 1). This design has two advantages over the original paradigm. First, measures of sensitivity in this experiment were based on the correct detection of a particular signal when it was present and the false detection of the same signal when it was not present. Second, the semantic consistency of the target label with the scene on corresponding target-present and catch trials was equivalent, because both labels named the same object. As a result, false-alarm rates in the semantically consistent condition were based entirely on the false detection of consistent target objects, and false-alarm rates in the semantically inconsistent condition were based entirely on the false detection of inconsistent target objects. In contrast to the original object-detection paradigm, this catch trial

design controls for the potential bias for participants to respond "yes" more often when the catch trial target label is consistent versus inconsistent with the scene.

To control for the general complexity of the scene in target-present and catch trials, the catch trial scenes contained the cued object from the paired scene at the cued location. For example, if a participant viewed the label "chicken" in a catch trial, the subsequent scene would contain the paired cued object (a mixer). Thus, in the catch trials, the semantic consistency manipulation was based on the relationship between the target label and the scene rather than the relationship between the cued object and the scene.

Method

Participants. Twenty-four Michigan State University undergraduate students participated in the experiment for course credit. All participants had normal or corrected-to-normal vision. The participants were naive with respect to the hypotheses under investigation. None had participated in Experiment 1.

Stimuli. The stimuli were the same as in Experiment 1 with the following modifications. First, for each scene, target labels in catch trials named the same object as in corresponding target-present trials. Second, the catch trial scenes contained the paired cued object (e.g., a mixer when the target object was a chicken).

Apparatus and procedure. The apparatus and procedure were the same as in Experiment 1. Each participant saw 160 experimental trials that were produced by a within-participant factorial combination of 2 target label semantic consistency conditions × 2 target object presence conditions × 2 cued object positions × 20 scenes. Because the cued object position factor was not of theoretical interest, the two levels of that factor were combined in the statistical analyses. Each participant saw all 160 trials in a different random order. The entire session lasted approximately 45 min.

Results

Percentage correct analysis. Mean percentage correct performance as a function of target label semantic consistency and target object presence is shown in Table 3. First, there was a reliable main effect of target object presence, $F(1, 23) = 8.61, MSE = .0739, p < .01$. Participants responded correctly 65.6% of the time when the target object was present and 77.1% of the time when the target object was absent. There was no main effect of target label semantic consistency, $F < 1$, with 71.9% correct performance when the target label was consistent with the scene and 70.9% when the target label was inconsistent. There was, however, a reliable interaction between target label semantic consistency and target object presence, $F(1, 23) = 65.85, MSE = .0269, p < .001$. The hit rate in the consistent target label condition (75.7%) was higher than that in the inconsistent target label condition (55.5%), but the reverse pattern obtained in the catch trials, with a higher correct rejection rate in the inconsistent target label condition (86.3%) than in the consistent target label condition (68.0%). In other words, both the hit and false-alarm rates were higher for consistent versus inconsistent target object-detection,

suggesting that participants were more biased to respond "yes" when attempting to detect a consistent target object. Given this pattern, sensitivity measures provide a better indication of detection accuracy than percentage correct performance.

A' analysis. Mean *A'* for each target object consistency condition is presented in Table 3. No effect of target label semantic consistency was obtained, *F* < 1. Participants were equally accurate at detecting semantically consistent target objects (.803) as semantically inconsistent target objects (.810).

Discussion

In Experiment 2, the catch trial design of the original object-detection paradigm was modified so that participants attempted to detect the same target object on corresponding target-present and catch trials. The main result was that no advantage was found for the detection of semantically consistent versus inconsistent target objects. In addition, there were reliable response biases caused by target label semantic consistency. As in Experiment 1, participants were biased to respond "yes" more often when the catch trial label was consistent versus inconsistent with the scene. This bias suggests that information sufficient to access scene meaning was available within the 200 ms scene presentation duration.

Why might the consistent object contextual facilitation effect, present in Experiment 1, be absent in Experiment 2? One potential explanation is that, in the catch trials, the presence of the paired cued object at the cued location may have biased performance. For example, in an inconsistent condition catch trial, the target label was "chicken"; this label was followed by a kitchen scene containing the paired cued object (a mixer). If consistent objects are easier to detect than inconsistent objects, as suggested by contextual facilitation models, false-alarm rates may have been artificially low when the target label was inconsistent with the scene, because on those trials, the cued object was always consistent with the scene. This lower false-alarm rate in the inconsistent condition might have masked consistent object facilitation in our sensitivity measure. We have no direct way to determine whether cued object semantic consistency influenced performance in this experiment, as that factor was confounded with the semantic consistency of the target label. The results from Experiment 1, as well as those reported by Biederman et al. (1982), however, provide no support for this hypothesis. Both experiments strongly

Table 3
Mean Percentage Hits, Percentage Correct Rejections (Percentage False Alarms), and A' for Experiment 2

Target label consistency	% hits	% correct rejections (% false alarms)	A'
Consistent	75.7	68.0 (32.0)	.803
Inconsistent	55.5	86.3 (13.7)	.810

suggest that the semantic consistency of the cued object has little if any effect on performance in catch trials, because false-alarm rates did not differ as a function of cued object consistency.

A second explanation for the difference in results between Experiments 1 and 2 is that the facilitated detection of consistent objects in Experiment 1 and in the original object-detection paradigm was an artifact produced by the method of calculating sensitivity. In these experiments, measures of sensitivity did not control response biases caused by the semantic consistency of the target label, as those paradigms averaged across catch trials on which the target object was consistent and inconsistent with the scene. In Experiment 2, when the catch trials were modified so that response biases were controlled in *A'*, no such consistent object advantage was obtained. This suggests that results from the original object-detection paradigm reflected response bias rather than the influence of scene context on object perception, and thus cannot be taken as strong evidence for contextual facilitation of consistent object perception.

A final explanation for the absence of a consistent object facilitation effect in Experiment 2 hinges on the use of a target label preview. The priming model of scene context effects proposes that the source of contextual facilitation effects is the spread of activation from the activated representation of a scene to stored descriptions of object types likely to be found in the scene (Friedman, 1979; Kosslyn, 1994; Palmer, 1975a). In Experiment 2, the presentation of the target label prior to scene viewing may have served to prime the stored description of the target object, regardless of its semantic consistency. Such priming could mask potential influences of scene context. In Experiment 3, the target label was presented either before or after the scene. If the priming model is correct, contextual facilitation of consistent object detection should be observed when the target label is presented after the scene, but not necessarily when the target label is presented before the scene.

Experiment 3

Experiment 3 sought to provide further evidence concerning the influence of scene context on object perception. In this experiment, the target label was presented either before or after scene presentation. In addition, the location cue was eliminated. Otherwise, Experiment 3 was identical to Experiment 2. The presentation of the target label after the scene further refines the object-detection paradigm by eliminating two potential problems. First, participants may have used the target label preview to constrain the spatial extent of their search in the subsequent scene. This strategy would benefit the detection of consistent objects, because the position of a consistent object in a scene is easier to predict than the position of an inconsistent object. Second, the presentation of a semantically inconsistent target label before the scene may have interfered with perceptual processing when the scene implied by the label was not presented. In addition, elimination of the location cue improves the object-
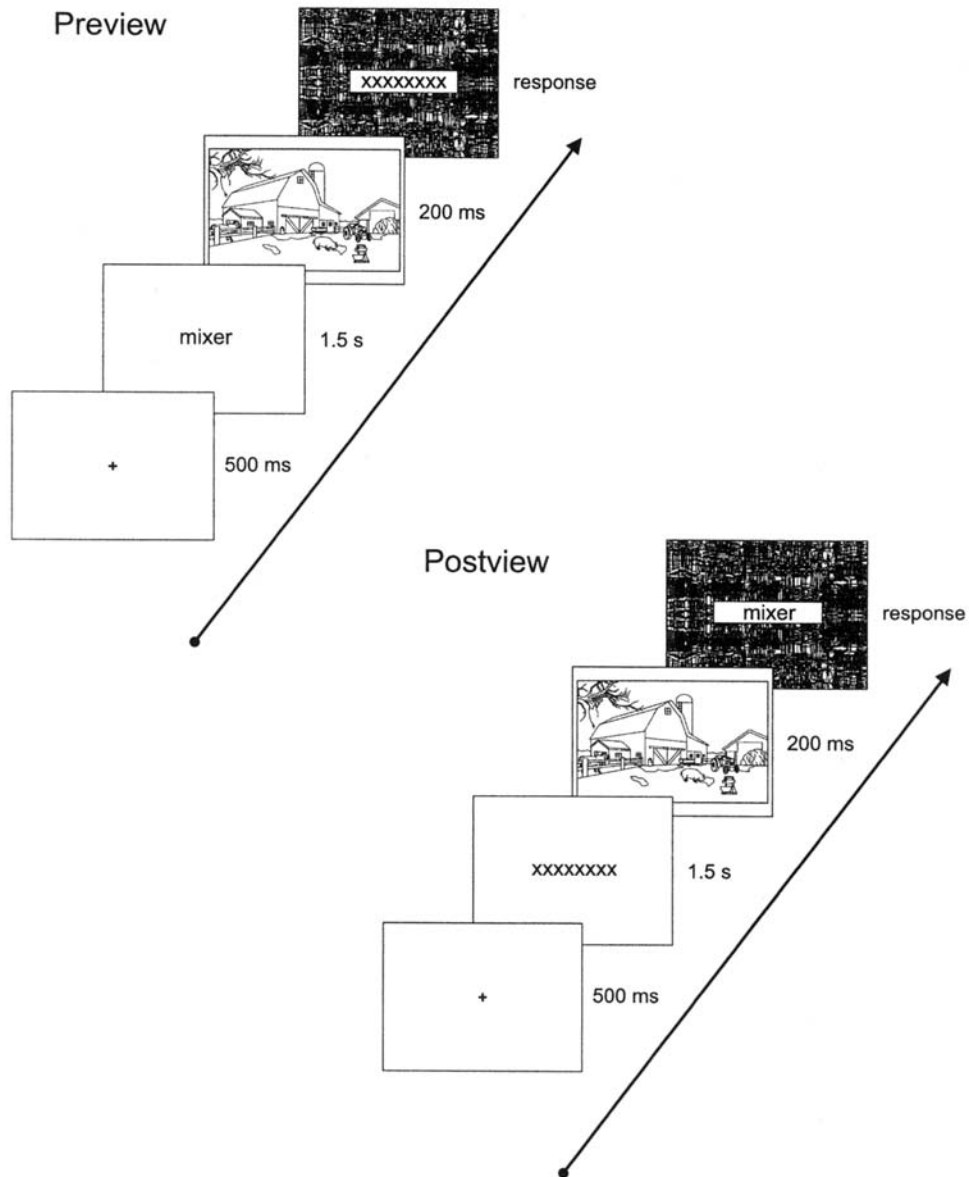
Figure 3. Schematic illustration of a trial in Experiment 3.

detection paradigm by undermining the potential strategy of using cue position to assist post-presentation guessing.

Figure 3 illustrates the main aspects of the paradigm. In the target label preview condition, participants saw a target label for 1500 ms, followed by presentation of the scene for 200 ms, followed by a pattern mask containing a series of lowercase *X*s, which remained on the screen until response. In the target label postview condition, participants saw a series of *X*s in a blank field for 1500 ms, followed by presentation of the scene for 200 ms, followed by a pattern mask containing the target label, which remained on the screen until response. The series of *X*s was employed to roughly equate stimulus presentation in the two conditions. In addition to the manipulation of target label (preview or postview), we omitted the location cue. The absence of a

location cue changed the participants' task. In Experiments 1 and 2, the task was to determine whether the target object appeared at the cued location. In Experiment 3, the task was to determine whether the target object appeared anywhere in the scene. As in Experiment 2, the semantic consistency manipulation in the catch trials was based on the relationship between the target label and the scene rather than the relationship between the presented object and the scene.[6] Table 1 presents a summary of the target-present and catch trial design in Experiment 3.

[6] In Experiments 3 and 4, the object presented in the scene was not cued by a location dot. Therefore, we will refer to this object as the *presented* object rather than the *cued* object.

Method

Participants. Twenty-four members of the Michigan State University community were paid $5 each for their participation. All participants had normal or corrected-to-normal vision. The participants were naive with respect to the hypotheses under investigation. None had participated in previous experiments.

Stimuli. The stimuli were the same as in Experiment 2, except that a blank space, subtending 6.3 × 1.5 degrees, was created in the center of the pattern mask to accommodate the object label or series of Xs.

Apparatus. The apparatus was the same as that used for Experiment 1, except that the stimuli were displayed on a flat-screen monitor with a 72-Hz refresh rate.

Procedure. The procedure was the same as that in Experiment 1, except that the participant was informed that the target label could appear either before or after the scene. In either case, the participant was to press the button marked "yes" if the object named by the label had appeared anywhere in the scene or the button marked "no" if the object had not appeared in the scene.

Each participant saw 160 experimental trials that were produced by a within-participant factorial combination of 2 target label presentation conditions (preview, postview) × 2 target label consistency conditions × 2 target object presence conditions × 20 scenes. Object position was manipulated between participants and was tied to the preview-postview manipulation. In one group, if position A was employed for the preview trials of a particular scene, position B was used for all postview trials employing that scene. These assignments were reversed in the second group. Because the object position factor was not of theoretical interest, the two levels of that factor were combined in the statistical analyses. Each participant saw all 160 trials in a different random order. The entire session lasted approximately 45 min.

Results

Percentage correct analysis. Mean percentage correct performance as a function of target label presentation, target label semantic consistency, and target object presence is shown in Table 4. First, there was a reliable main effect of target label presentation (preview, postview), $F(1, 23) = 36.00$, $MSE = .0082$, $p < .001$. Participants responded correctly 73.8% of the time in the preview condition and 65.9% of the time in the postview condition. There was also a main effect of target object presence, $F(1, 23) = 10.89$, $MSE = .0765$, $p < .005$, with better performance in the catch trials (76.4%) than in the target-present trials (63.2%). There was no main effect of target label semantic consistency, $F < 1$, but there was a reliable interaction between target label semantic consistency and target object presence, $F(1, 23) = 107.67$, $MSE = .0313$, $p < .001$. The hit rate for consistent target objects (76.4%) was higher than that for inconsistent target objects (50.1%), but the reverse pattern obtained in the catch trials, with a higher correct rejection rate for inconsistent target objects (89.8%) than for consistent target objects (63.0%). Finally, there was a reliable 3-way interaction between all factors considered, $F(1, 23) = 10.63$, $MSE = .0080$, $p < .005$. In the target-present trials, performance for consistent target objects dropped slightly between preview and postview conditions (6.1%), but performance for inconsistent target objects dropped more significantly (13.2%). In the catch trials, the reverse pattern

obtained: Correct rejections for consistent target objects dropped significantly between preview and postview trials (11.0%), but correct rejections for inconsistent target objects showed almost no decline (1.2%). As in Experiments 1 and 2, both the hit rate and false-alarm rate were higher for consistent versus inconsistent target objects, suggesting that participants were biased to respond "yes" more often when attempting to detect a consistent versus an inconsistent object. Given this pattern, sensitivity measures again provide a better indication of detection accuracy than percentage correct performance.

A' *analysis.* Mean A' as a function of target label presentation and target label semantic consistency is reported in Table 4. There was a reliable main effect of target label presentation, $F(1, 23) = 23.52$, $MSE = .0067$, $p < .001$, with better detection in the preview conditions (.836) than in the postview conditions (.755). In addition, there was a reliable main effect of semantic consistency, $F(1, 23) = 5.87$, $MSE = .0034$, $p < .05$, with better detection of semantically inconsistent target objects (.810) than consistent target objects (.782). These effects were mediated by a marginally reliable interaction between label presentation and target object semantic consistency, $F(1, 23) = 3.55$, $MSE = .0036$, $p = .07$. Planned simple effects tests showed that there was no difference in performance for consistent versus inconsistent target objects in the preview condition, $F < 1$, but there was a reliable advantage for inconsistent target object detection in the postview condition, $F(1, 23) = 5.35$, $MSE = .0060$, $p < .05$.

Discussion

The results from the preview condition of Experiment 3 replicated those of Experiment 2 and provided no support for contextual facilitation models of object perception in scenes. The pattern of performance as a function of target label consistency was the same across the two experiments: There was no advantage for the detection of semantically consistent versus inconsistent objects, and participants demonstrated a bias to respond that consistent target objects were present in the scene compared to inconsistent objects. In addition, overall performance in the preview condition of Experiment 3 ($A' = .836$) was higher than that in Experiment 2 ($A' = .807$). These experiments differed only in the presence of the location cue, suggesting that the location cue provided little if any information to aid detection perfor-

Table 4
Mean Percentage Hits, Percentage Correct Rejections (Percentage False Alarms), and A' for Experiment 3

Target label consistency	% hits	% correct rejections (% false alarms)	A'
Preview			
Consistent	79.4	68.5 (31.5)	.834
Inconsistent	56.7	90.4 (9.6)	.839
Postview			
Consistent	73.3	57.5 (42.5)	.729
Inconsistent	43.5	89.2 (10.8)	.781

mance. The results of the postview condition of Experiment 3 also failed to support contextual facilitation models of object perception in scenes. Contrary to the prediction of those models, an advantage for the detection of inconsistent target objects was obtained. This result is in particular contrast to the prediction derived from the priming model that consistent object facilitation may only be observed when the target label is presented after the scene.

In addition, the results from Experiment 3 provide further evidence that the semantic consistency between the cued object and the scene has little or no influence on catch trial performance. In Experiment 3, the location cue was eliminated, and the task was to determine whether the target object appeared anywhere in the scene. Participants were therefore unaware, at least initially, that the identity of the presented object had any bearing on whether the target object was present or absent. If the lower false-alarm rate for inconsistent versus consistent target objects in Experiment 2 was caused by the semantic consistency of the cued object, that difference should have disappeared, or at least have been attenuated, in Experiment 3. Contrary to this prediction, the false-alarm rates for inconsistent target label catch trials in both the preview and postview conditions of Experiment 3 were lower than that in Experiment 2. In addition, the difference in the false-alarm rate as a function of target label semantic consistency was actually larger in both the preview and postview conditions of Experiment 3 than in Experiment 2, suggesting that participants were more biased to respond that consistent objects were present in Experiment 3 than in Experiment 2. Thus, the absence of a consistent object advantage in Experiment 2 and the preview condition of Experiment 3, and the presence of the inconsistent object advantage in the postview condition of Experiment 3, do not appear to have been caused by the semantic consistency of the object presented in the scene during catch trials.

Experiment 4

The purpose of Experiment 4 was to provide converging evidence bearing on the general hypothesis that consistent scene context facilitates object perception. In Experiments 1–3, A' was employed to control for participant response biases. In Experiment 4, we introduced a forced-choice procedure, similar to that developed by Reicher in the word recognition literature (Reicher, 1969), to eliminate response bias entirely. This paradigm eliminates response biases caused by the semantic consistency of the object label because participants must discriminate between two object labels, both of which are either semantically consistent or inconsistent with the scene. Figure 4 illustrates the main aspects of the paradigm.

A scene was presented for 250 ms, followed by a pattern mask for 30 ms, followed by a screen displaying two object labels. One label named an object presented in the scene, and the other label named an object that had not appeared in the scene. The participants' task was to indicate which of the two labels named an object that had been presented in the scene. The main contextual manipulation was the semantic

consistency between the scene and the presented object. For this experiment we chose one additional consistent and one additional inconsistent object to be presented in each scene. Thus, each scene could contain one of four presented objects, two of which were consistent and two of which were inconsistent. In the forced-choice response screen, one object label named the object presented in the scene, and the second label named the other object of the equivalent semantic consistency. For example, when a consistent object (a chicken or a pig) was presented in the farmyard scene, the forced choice screen presented the labels "chicken" and "pig." As in previous experiments, contextual facilitation models predict that percentage correct discrimination performance should be better when the presented object is consistent versus inconsistent with the scene in which it appears.

Method

Participants. Twenty-four Michigan State University undergraduate students participated in the experiment for course credit. One of these original participants had to be replaced because he had difficulty understanding the instructions. All participants had normal or corrected-to-normal vision. The participants were naive with respect to the hypotheses under investigation. None had participated in previous experiments.

Stimuli. The stimuli were the same as in Experiments 1–3 with the following modifications. For each scene, a second consistent and a second inconsistent object were chosen. One scene from the original set was replaced because of the difficulty of finding a second object that would be consistent in the scene at the same location as the original consistent object. Finally, only one object location was used for each scene. The object labels appearing after scene presentation were centered vertically and positioned to the left and right of fixation. The labels were created using lower-case, 24-point, anti-aliased Arial font.

Apparatus and procedure. The apparatus was the same as in Experiments 1 and 2. Participants were presented a scene for 250 ms, followed by a pattern mask for 30 ms, followed by a forced-choice screen containing two object labels. There was no delay (i.e., the inter-stimulus interval was zero) between each display. The forced-choice screen remained in view until the participant pressed the left button to indicate that the object named by the left-hand label had appeared or the right button to indicate that the object named by the right-hand label had appeared in the scene. Each participant saw 160 experimental trials that were produced by a within-participant factorial combination of 2 presented object consistency conditions × 2 presented objects × 2 label positions in the forced-choice display × 20 scenes. Because the presented object factor and the label position factor were not of theoretical interest, the two levels of each factor were combined in the statistical analyses. Each participant saw all 160 trials in a different random order. The entire session lasted approximately 45 min.

Results

The influence of object consistency on percentage correct discrimination performance was analyzed via a simple effects test. There was no effect of the consistency of the presented object, $F < 1$. Participants responded correctly 70.7% of the time when the presented object was consistent
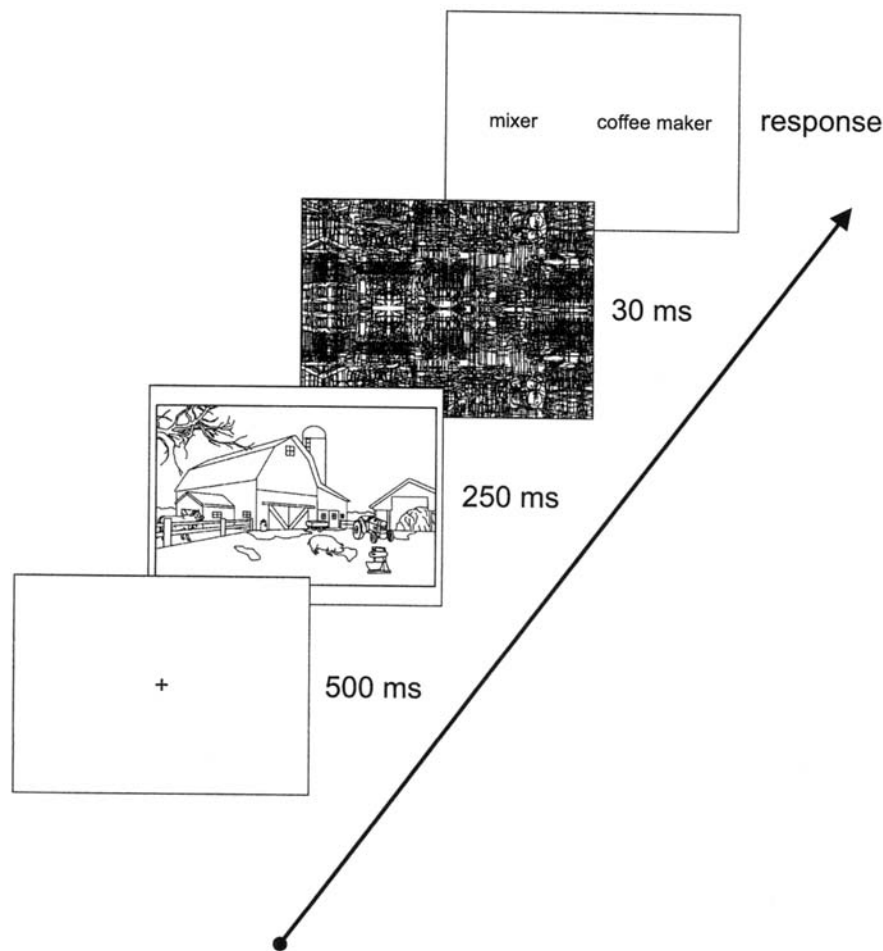
Figure 4. Schematic illustration of a trial in Experiment 4.

with the scene and 71.6% of the time when the presented object was inconsistent with the scene. The 95% confidence interval around these means was ±1.69%. Thus, the experiment had enough power to detect a 2.39% effect (see Loftus & Masson, 1994).

Discussion

In Experiment 4, we introduced a forced-choice procedure to investigate the influence of scene context on object perception independently of response bias. Contrary to the prediction of contextual facilitation models, no advantage for the detection of consistent objects was found. In fact, the non-reliable trend was in the direction of better inconsistent object detection. Thus, using the same general set of scene stimuli, the original object-detection paradigm (Experiment 1) produced a consistent object advantage, but Experiment 4, in which response bias was eliminated, showed no such advantage. This suggests, again, that the consistent object advantage in the original object-detection paradigm resulted, at least in part, from the inadequate control of response bias rather than from the influence of scene context on object perception.

In Experiment 4, only one object position was used for each scene. It is possible that the absence of a consistent object advantage was due to participants learning the position at which the presented object appeared in each scene. Such knowledge could allow participants to direct their attention quickly to the object position regardless of whether the presented object was semantically consistent or inconsistent with the scene. To investigate this possibility, we calculated percentage correct discrimination performance as a function of object consistency and first half versus second half of the trials. If the learning of object positions masked a consistent object advantage, that advantage would more likely be found in the first half of the experiment than in the second. However, there was no evidence of a consistent object advantage in the first half of the experiment. Percentage correct discrimination was 67.3% when the presented object was consistent and 68.8% when the presented object was inconsistent. A second potential concern with Experiment 4 is that the scene was presented for 250 ms, 100 ms longer than studies finding contextual facilitation of object detection (Biederman et al., 1982; Boyce et al., 1989). However, we have replicated Experi-

ment 4 with a scene presentation duration of 150 ms and obtained no advantage for the discrimination of semantically consistent objects (Hollingworth & Henderson, in press-a). In fact, discrimination performance was reliably higher when the target object was inconsistent versus consistent with the scene.

General Discussion

The purpose of this study was to investigate whether object identification is influenced by the semantic relationship between an object and the scene in which it appears. One view of the influence of scene context on object identification proposes that consistent scene context facilitates the perception of objects (Biederman, 1972; Biederman et al., 1973; Kosslyn, 1994; Ullman, 1996). This view has been instantiated in contextual facilitation models of object perception in scenes, which propose that knowledge about the objects found in a given scene type facilitates the perception of object stimuli consistent with that scene (Biederman, 1981; Biederman et al., 1982; Boyce & Pollatsek, 1992; Boyce et al., 1989; Friedman, 1979; Palmer, 1975a). The primary evidence supporting contextual facilitation models comes from experiments demonstrating that the detection of an object in a scene is facilitated when the object is semantically consistent compared with when it is inconsistent with the scene (Biederman et al., 1982; Boyce et al., 1989). However, as discussed in the Introduction, there are several potential problems with the original object-detection paradigm that make interpretation of these data difficult (De Graef et al., 1990; De Graef & d'Ydewalle, 1995; Henderson, 1992). Specifically, the original paradigm may not have adequately controlled for participant response biases, and may have provided additional sources of information that influenced detection performance. In this study, we modified the original object-detection paradigm to address these concerns.

In Experiment 1, we replicated the original object-detection paradigm contextual facilitation effect (Biederman et al., 1982; Boyce et al., 1989). Participants saw a target label naming an object, followed by a briefly presented scene, followed by a pattern mask with an embedded location cue. The object marked by the location cue could be either semantically consistent or inconsistent with the scene in which it appeared. Detection performance was based on percentage correct for trials in which the target label named the cued object and percentage of errors for trials in which the target label named a different object that did not appear in the scene. In these latter trials, the semantic consistency of the target label with the subsequent scene was not related to the consistency of the cued object with the scene. Using this paradigm, we replicated the basic consistent object contextual facilitation effect: Detection performance (A') was better when the cued object was semantically consistent versus inconsistent with the scene. However, a more detailed analysis of the catch trial data indicated that reliable response biases were caused by the semantic consistency of the target label: Participants responded "yes" more often when the catch trial target label was semantically consistent

versus inconsistent with the scene. Because of the design of the catch trials in this paradigm, sensitivity measures were unable to adequately control for this response bias.

In Experiment 2, we modified the original object-detection paradigm so that participants attempted to detect the same object on trials in which it was and was not present in the scene. This modification assured that the detection performance measure (A') was based on the correct detection of a particular signal when it was present in the scene and the false detection of the same signal when it was not present. In addition, response biases caused by the semantic consistency of the target label were controlled in A', providing a valid measure of object-detection performance. Contrary to the prediction of contextual facilitation models, no advantage was found for the detection of consistent versus inconsistent objects.

In Experiment 3, we manipulated whether the target label appeared before or after the scene, and we eliminated the location cue. The purpose of the label manipulation was to investigate whether presenting the label before the scene may have interfered with scene processing when it was inconsistent with the subsequent scene, and may have allowed participants to constrain their subsequent search, biasing detection performance. The location cue was eliminated to see if it provided information useful to post-presentation guessing strategies. The results from the target label preview condition replicated those of Experiment 2 and indicated that the location cue provided little or no information to aid performance. The main result from the target label postview condition was superior detection of inconsistent versus consistent objects.

In Experiment 4, we employed a forced-choice procedure to assess object-detection performance. This procedure improved the object-detection paradigm by eliminating the possibility of response bias caused by the semantic consistency of the target label. Participants were asked to discriminate between two consistent object labels (when the presented object was consistent with the scene) or between two inconsistent object labels (when the presented object was inconsistent with the scene). Contrary to the prediction of contextual facilitation models, no advantage was found for the detection of consistent objects.

In order to conclude from these data that consistent scene context does not facilitate object identification, it is necessary that these experiments meet two criteria. First, scene meaning must have been available early enough to influence object identification, if such influences exist. Second, the contextual manipulation must have been strong enough to interact with object identification, if such interaction is possible. In this study, scene meaning was available early enough and was strong enough to produce reliable response biases based on the consistency of the target label. These response biases suggest that scene meaning and its relationship to the identity of the target object was available from information obtained within the brief presentation duration of the scene stimulus. In addition, the contextual manipulation was strong enough to replicate the Biederman et al. (1982) results. Given that we found no evidence for facilitated detection of semantically consistent objects in the

face of a robust contextual manipulation, the conclusion follows that consistent scene context does not facilitate object identification.

In summary, when the object-detection paradigm was modified so that (a) response biases were either adequately controlled or eliminated from the paradigm and (b) other sources of potentially biasing information were removed, no advantage was found for the detection of semantically consistent objects. This suggests that the results from previous object-detection paradigms (Biederman et al., 1982; Boyce et al., 1989) may not have reflected the influence of scene context on object identification. Instead, the consistent object advantage in these experiments appears to have been caused by the inadequate control of response bias and by the presentation of the target label before the scene. Most significantly, these data suggest that consistent scene context does not facilitate the identification of real-world objects.

Interactivity Versus Functional Isolation

The data from this study do not support the hypothesis that stored knowledge about a scene type facilitates the identification of consistent object stimuli, as proposed by contextual facilitation models of object perception in scenes. In contrast, these data are consistent with a functional architecture of the visual system in which object perception is isolated from stored information about the objects likely to appear in a scene. Such isolation may be necessary to avoid what has been termed the *frame problem* (Fodor, 1983). Given that experience with natural scenes comprises most of waking life, the set of potentially relevant information to any scene perception task would be so large as to make the discrimination between relevant and irrelevant information resource intensive and time consuming. Therefore, perceptual systems may be fast and accurate precisely because they do not consult sources of information such as contextually derived expectations.

Alternatively, it might be argued that functional isolation derives not from structural properties of the visual system but from the relatively low degree of constraint that exists in the real world between scenes and objects. By way of illustration, consider the relationship between letters and words versus objects and scenes. A particular word is present if and only if a certain set of letters is present in a particular spatial arrangement. In contrast, the constraints between a given scene and the objects in that scene are far less strict. For example, a kitchen scene remains a kitchen scene even if a highly diagnostic object, such as a stove, is not present. In addition, the spatial arrangement of the objects in a scene is far less constrained; as long as basic physical constraints are upheld, a stove can appear in many different locations relative to other objects in the scene. Moreover, scene and object representations are not necessarily organized hierarchically; whereas a word must be recognized through local analysis of letter identities, a scene can be recognized when most of the local object information is removed via low-pass filter (e.g., Schyns & Oliva, 1994). Thus, it might be argued that even if conceptual information can, in principle, interact

with perceptual systems, in the case of object perception, functional isolation between object perception and scene knowledge has been produced by the relatively weak constraints that a given scene places on the objects and spatial relationships likely to be found there.[7]

This hypothesis, though logically possible, is not very satisfying. First, although the correlations between specific scenes and specific objects are not one, they are also not zero. In the case of the stimuli used in the present experiments, the target objects were relatively highly semantically constrained. For example, given that one is looking at a farm scene, a chicken is relatively likely to appear, whereas a mixer is relatively unlikely. This type of relationship does not seem to be unusual for real-world scenes. More generally, the view that less than perfect correlations in the world can lead to functional isolation seems to run counter to current perspectives on constraint-based perception, which assume that all available constraints are consulted when interpreting an input pattern (Mumford, 1994; Rumelhart et al., 1986).

Although the results of this study are consistent with the hypothesis that object perception is functionally isolated from scene knowledge, they may appear to be contrary to evidence that suggests facilitative interaction between other visual subsystems (where a sub-system is optimized to perform a specific task). For example, classic demonstrations of the effects of perceptual learning on difficult segmentation problems, such as that seen with the Dalmatian dog image (Gregory, 1970; Neisser, 1967), suggest that the presence of a stored image interpretation can facilitate initial segmentation and grouping processes for that image.[8] However, these results are not incompatible with ours. Facilitative interaction (or functional isolation) between two visual subsystems cannot be taken as evidence that such interaction (or functional isolation) exists between all subsystems. There may be functional isolation between some forms of representation (e.g., scene conceptual knowledge and object perception) but not between others (e.g., stored object models and segmentation routines). More generally, in order to investigate interactivity and isolation in the visual system, it is important to specify explicitly the representational systems that are potentially functional in a given

---

[7] The fact that the constraints are weak between scenes and objects might lead to a functional isolation of object perception either ontogenetically or phylogenetically. The former case would be consistent with constraint-satisfaction theories of cognition, in which the contingencies experienced by a specific individual determines the functional architecture of that individual's visual-cognitive system. Alternatively, those same constraints, experienced over evolutionary time scales, may have produced a genetically programmed neural architecture in which functional isolation between object perception and scene knowledge is produced.

[8] More recently, Peterson and Gibson (1994) have provided evidence that figure-ground segmentation for an object is not independent of possible object interpretations. Results from these experiments might also be interpreted as indicating that there is interaction between stored object models and the computation of a perceptual description of an object, though it should be noted that Peterson and Gibson do not interpret their results in this way.

visual task, to determine what interactions are logically possible among these systems, and then to examine empirically each potential interaction.

This point is also relevant to recent theorizing about the functional role of feedback projections in the neural architecture of the visual system. It is tempting to conclude that because there are back-projecting pathways from later to earlier cortical visual areas (e.g., Desimone & Ungerleider, 1989), there must be massive top-down influences at all levels of visual analysis, including an influence of scene meaning on the generation of a description of input patterns for objects (e.g., Ullman, 1996). However, it is premature to make this inference given our current understanding of the relationship between neural architecture and functional architecture. The existence of re-entrant neural pathways does not necessarily imply the type of massive interactivity between sub-systems assumed by the hypothesis that vision is a constraint-satisfaction problem in which all sub-systems settle together (e.g., Mumford, 1994). For example, re-entrant pathways could be serving one or more alternative functions, such as attentional regulation of information flow between cortical areas (Posner & Petersen, 1990; Van Essen, Anderson, & Olshausen, 1994) or the binding of visual representations to other visual representations, spatial representations, and motor representations (Henderson, 1996; Tanaka, 1996; Van Essen et al., 1994). Thus, the existence of re-entrant neural pathways is not necessarily evidence for massive constraint-satisfaction interactivity between functional sub-systems.

The Inconsistent Object Advantage

There is some evidence in this study to indicate that detection performance was not entirely insensitive to the semantic relationship between an object and the scene in which it appeared. In the postview condition of Experiment 3, detection performance was better for semantically inconsistent versus consistent objects. Although this result may appear anomalous, it replicates an inconsistent object advantage found in a similar paradigm based on change detection (Henderson, Weeks, & Hollingworth, 1996; Hollingworth & Henderson, in press-b). In this paradigm, the participant was presented with an initial picture of a natural scene for 250 ms, followed by a pattern mask, followed by presentation of a test scene. The test scene either was identical to the initial scene or was identical except for a change to a single target object in the scene (either deletion or mirror reversal). The participant's task was to determine whether any of the objects in the scene had changed across the two presentations. The semantic consistency of the target object was manipulated, so that it was either semantically consistent or inconsistent with the scene in which it appeared. Detection of changes, as reflected in A', was better when the object undergoing the change was semantically inconsistent with the scene.

It is possible that this inconsistent object advantage could be explained by facilitatory interactions between stored scene knowledge and the perceptual processing of objects inconsistent with the scene. This idea seems unlikely, however,

because it assumes that a scene concept represents information about each object in the extremely large set of objects that would be inconsistent with that scene. Thus, we propose two potential hypotheses to explain the inconsistent object advantage, both of which are compatible with the view that object perception is functionally isolated from stored knowledge about the objects likely to be found in the scene. These are the *memory schema hypothesis* and the *attention hypothesis*.

The memory schema hypothesis. According to this hypothesis, perceptual encoding of semantically consistent and inconsistent objects proceeds equivalently, but following perceptual encoding, information about semantically inconsistent objects is preferentially remembered. Specifically, information about objects consistent with the scene may be lost during a normalization process within a memory schema, whereas information about inconsistent objects may be retained more veridically, perhaps as part of a list noting deviations from the default values in the schema (Friedman, 1979). This hypothesis is supported by scene memory studies that have shown better long-term memory for semantically inconsistent versus consistent objects (e.g., Friedman, 1979). In Experiment 3, the object label was presented immediately after the offset of the scene. Thus, if the inconsistent object advantage in Experiment 3 is due to memory schema effects, these processes must occur quite rapidly.

The attention hypothesis. According to the attention hypothesis, perceptual encoding of semantically consistent and inconsistent objects is not influenced directly by scene context, but, after an object has been identified, attention is preferentially allocated to objects that violate the constraints imposed by scene meaning. Attention may be allocated to inconsistent objects because they are difficult to integrate into the conceptual representation established by the scene and require the encoding of more detailed perceptual information to resolve the contextual discrepancy. The additional attentional resources devoted to an inconsistent object would then produce a more complete perceptual description of that object, leading to better detection performance in Experiment 3.

Support for the attention hypothesis comes from eye movement studies that have measured the allocation of overt attention in scenes containing semantically consistent and inconsistent objects. A number of studies have indicated that once an inconsistent object in a scene is fixated, the eyes tend to dwell longer on that object compared to consistent objects (De Graef et al., 1990; Friedman, 1979; Henderson et al., in press; Loftus & Mackworth, 1978). In addition, Loftus and Mackworth (1978) have provided evidence that the eyes may be drawn to inconsistent objects in the periphery, though this result has not been replicated with realistic scene stimuli such as those used in the present study (Henderson et al., in press; De Graef et al., 1990; see Henderson & Hollingworth, 1998). Thus, eye movement studies suggest that although attention may not be initially drawn to regions of semantic inconsistency, once such a region has been attended, attention may be captured and may dwell longer at that location. This longer attentional dwell

time, then, could increase the quality of the information encoded for a semantically inconsistent object.

Conclusion

In summary, we reported evidence from four experiments showing that when the original object-detection paradigm was modified to control for participant response biases, and when other sources of potentially biasing information were eliminated, no advantage was found for the detection of semantically consistent versus inconsistent objects. These results do not support contextual facilitation models of object identification in scenes (Biederman, 1981; Biederman et al., 1982; Boyce & Pollatsek, 1992; Boyce et al., 1989; Friedman, 1979; Palmer, 1975a), nor do they support the general hypothesis that object perception is influenced by the constraints imposed by scene meaning (Biederman, 1972; Biederman et al., 1973; Kosslyn, 1994; Ullman, 1996). Instead, the results suggest that object identification processes may be functionally isolated from scene contextual information. Once an object representation has been formed, however, the semantic status of the object with the scene may influence memory for that object or may influence the allocation of visual/spatial attention, leading to an enhanced representation of objects inconsistent with the scene.

References

Antes, J. R., Penland, J. G., & Metzger, R. L. (1981). Processing global information in briefly presented scenes. *Psychological Research, 43*, 277–292.

Barlow (1994). What is the computational goal of the neocortex? In C. Koch & J. L. Davis (Eds.), *Large scale neuronal theories of the brain* (pp. 1–22). Cambridge, MA: MIT Press.

Biederman, I. (1972, July). Perceiving real-world scenes. *Science, 177*, 77–80.

Biederman, I. (1981). On the semantics of a glance at a scene. In M. Kubovy & J. R. Pomerantz (Eds.), *Perceptual organization* (pp. 213–253). Hillsdale, NJ: Erlbaum.

Biederman, I. (1987). Recognition-by-components: A theory of human image understanding. *Psychological Review, 94*, 115–147.

Biederman, I., Glass, A. L., & Stacy, E. W. (1973). Searching for objects in real-world scenes. *Journal of Experimental Psychology, 97*, 22–27.

Biederman, I., Mezzanotte, R. J., & Rabinowitz, J. C. (1982). Scene perception: Detecting and judging objects undergoing relational violations. *Cognitive Psychology, 14*, 143–177.

Boyce, S. J., & Pollatsek, A. (1992). Identification of objects in scenes: The role of scene background in object naming. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18*, 531–543.

Boyce, S. J., Pollatsek, A., & Rayner, K. (1989). Effect of background information on object identification. *Journal of Experimental Psychology: Human Perception and Performance, 15*, 556–566.

Bruner, J. S. (1957). On perceptual readiness. *Psychological Review, 64*, 123–152.

Bruner, J. S. (1973). *Beyond the information given.* New York: Norton.

Bülthoff, H. H., Edelman, S. Y., & Tarr, M. J. (1995). How are three-dimensional objects represented in the brain? *Cerebral Cortex, 3*, 247–260.

Churchland, P. S., Ramachandran, V. S., & Sejnowski, T. J. (1994). A critique of pure vision. In C. Koch and J. L. Davis (Eds.), *Large scale neuronal theories of the brain* (pp. 23–60). Cambridge, MA: MIT Press.

De Graef, P., Christiaens, D., & d'Ydewalle, G. (1990). Perceptual effect of scene context on object identification. *Psychological Research, 52*, 317–329.

De Graef, P., & d'Ydewalle, G. (1995). *Speeded object verification in real-world scenes: Perceptual, decisional, and attentional components* (Rep. No. 170). Leuven, Belgium: University of Leuven, Laboratory of Experimental Psychology.

Desimone, R., & Ungerleider, L. G. (1989). Neural mechanisms of visual processing in monkeys. In F. Boller & J. Grafman (Eds.), *Handbook of neuropsychology* (Vol. 2, pp. 267–299). New York: Elsevier.

Fodor, J. A. (1983). *Modularity of mind.* Cambridge, MA: MIT Press.

Friedman, A. (1979). Framing pictures: The role of knowledge in automatized encoding and memory for gist. *Journal of Experimental Psychology: General, 108*, 316–355.

Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics.* New York: Wiley.

Gregory, R. L. (1970). *The intelligent eye.* London: Weidenfield & Nicholson.

Grier, J. B. (1971). Nonparametric indexes for sensitivity and bias: Computing formulas. *Psychological Bulletin, 75*, 424–429.

Henderson, J. M. (1992). Object identification in context: The visual processing of natural scenes [Special issue]. *Canadian Journal of Psychology, 46*(3), 319–341.

Henderson, J. M. (1996). Visual attention and the attention–action interface. In K. A. Akins (Ed.), *Vancouver studies in cognitive science: Vol. 5. Perception* (pp. 290–316). Oxford, England: Oxford University Press.

Henderson, J. M., & Hollingworth, A. (1998). Eye movements during scene viewing: An overview. In G. Underwood (Ed.), *Eye guidance in reading and scene perception* (pp. 269–293). Oxford, England: Elsevier.

Henderson, J. M., & Hollingworth, A. (in press). High-level scene perception. *Annual Review of Psychology.*

Henderson, J. M., Weeks, P. A., Jr., & Hollingworth, A. (1996, November). *The influence of scene context on object perception.* Paper presented at the Annual Meeting of the Psychonomic Society, Chicago.

Henderson, J. M., Weeks, P. A., Jr., & Hollingworth, A. (in press). The effects of semantic consistency on eye movements during complex scene viewing. *Journal of Experimental Psychology: Human Perception and Performance.*

Hollingworth, A., & Henderson, J. M. (in press-a). Object identification is isolated from scene semantic constraint: Evidence from object type and token discrimination [Special issue]. *Acta Psychologica.*

Hollingworth, A., & Henderson, J. M. (in press-b). Semantic informativeness mediates the detection of changes in natural scenes [Special issue]. *Visual Cognition.*

Kosslyn, S. M. (1994). *Image and brain.* Cambridge, MA: MIT Press.

Loftus, G. R., & Mackworth, N. H. (1978). Cognitive determinants of fixation location during picture viewing. *Journal of Experimental Psychology: Human Perception and Performance, 4*, 565–572.

Loftus, G. R., & Masson, M. E. J. (1994). Using confidence intervals in within-subjects designs. *Psychonomic Bulletin & Review, 1,* 476–490.

Loftus, G. R., Nelson, W. W., & Kallman, H. J. (1983). Differential acquisition rates for different types of information from pictures. *Quarterly Journal of Experimental Psychology, 35A,* 187–198.

Mandler, J. M., & Johnson, N. S. (1976). Some of the thousand words a picture is worth. *Journal of Experimental Psychology: Human Learning and Memory, 2,* 529–540.

Mandler, J. M., & Parker, R. E. (1976). Memory for descriptive and spatial information in complex pictures. *Journal of Experimental Psychology: Human Learning and Memory, 2,* 38–48.

Marr, D. (1982). *Vision.* San Francisco, CA: Freeman.

Marr, D., & Nishihara, H. K. (1978). Representation and recognition of the spatial organization of three-dimensional shapes. *Proceedings of the Royal Society of London, B200,* 269–294.

McClelland, J. J., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: I. An account of basic findings. *Psychological Review, 88,* 375–407.

Metzger, R. L., & Antes, J. R. (1983). The nature of processing early in picture perception. *Psychological Research, 45,* 267–274.

Mumford, D. (1994). Neuronal architectures for pattern-theoretic problems. In C. Koch & J. L. Davis (Eds.), *Large scale neuronal theories of the brain* (pp. 125–152). Cambridge, MA: MIT Press.

Neisser, U. (1967). *Cognitive psychology.* Englewood Cliffs, NJ: Prentice Hall.

Palmer, S. E. (1975a). The effects of contextual scenes on the identification of objects. *Memory & Cognition, 3,* 519–526.

Palmer, S. E. (1975b). Visual perception and world knowledge: Notes on a model of sensory-cognitive interaction. In P. A. Norman, D. E. Rumelhart, & LNR Research Group (Eds.), *Explorations in cognition* (pp. 279–307). San Francisco: Freeman.

Peterson, M. A., & Gibson, B. S. (1994). Must figure–ground organization precede object recognition? An assumption in peril. *Psychological Science, 5,* 253–259

Posner, M. I., & Petersen, S. E. (1990). The attention system of the human brain. *Annual Review of Psychology, 13,* 25–42.

Potter, M. C. (1976). Short-term conceptual memory for pictures. *Journal of Experimental Psychology: Human Learning and Memory, 2,* 509–522.

Pylyshyn, Z. (1980). Computation and cognition: Issues in the foundations of cognitive science. *Behavioral and Brain Sciences, 3,* 111–132.

Reicher, G. M. (1969). Perceptual recognition as a function of meaningfulness of stimulus material. *Journal of Experimental Psychology, 81,* 275–280.

Rumelhart, D. E., McClelland, J. L., & the PDP Research Group. (1986). *Parallel distributed processing: Explorations in the microstructure of cognition: Vol. 1. Foundations.* Cambridge, MA: MIT Press.

Schyns, P. G., & Oliva, A. (1994). From blobs to boundary edges: Evidence for time- and spatial-scale-dependent scene recognition. *Psychological Science, 5,* 195–200.

Tanaka, K. (1996). Inferotemporal cortex and object vision. *Annual Review of Neuroscience, 19,* 109–139.

Ullman, S. (1996). *High-level vision: Object recognition and visual cognition.* Cambridge, MA: MIT Press.

van Diepen, P. M. J., & De Graef, P. (1994). *Line-drawing library and software toolbox* (Rep. No. 165). Leuven, Belgium: University of Leuven, Laboratory of Experimental Psychology.

Van Essen, D. C., Anderson, C. H., & Olshausen, B. A. (1994). Dynamic routing strategies in sensory, motor, and cognitive processing. In C. Koch & J. L. Davis (Eds.), *Large scale neuronal theories of the brain* (pp. 271–299). Cambridge, MA: MIT Press.