Memory & Cognition
2007, 35 (3), 393-401

Consistency effects between objects in scenes

Jodi L. Davenport
Massachusetts Institute of Technology, Cambridge, Massachusetts

How does context influence the perception of objects in scenes? Objects appear in a given setting with surrounding objects. Do objects in scenes exert contextual influences on each other? Do these influences interact with background consistency? In three experiments, we investigated the role of object-to-object context on object and scene perception. Objects (Experiments 1 and 3) and backgrounds (Experiment 2) were reported more accurately when the objects and their settings were consistent than when they were inconsistent, regardless of the number of foreground objects. In Experiment 3, related objects (from the same setting) were reported more accurately than were unrelated objects (from different settings), independently of consistency with the background. Consistent with an interactive model of scene processing, both object-to-object context and object–background context affect object perception.

How does context influence the perception of objects and scenes? Although recognition experiments typically investigate objects in isolation, objects in the world rarely appear without some context. Objects are always located spatially within a setting and usually appear with other objects. Does contextual information affect recognition? For example, a new acquaintance may be recognized more quickly in the presence of a mutual friend than alone. Here, in three experiments, we investigated whether and how objects in scenes influence the perception of each other and their background when viewing time is brief. The experiments bridge two lines of research on semantic consistency effects: (1) the influence of background context on perception of objects in scenes and (2) the influence of object relatedness outside of scenes.

Influence of Background Context on the Perception of Objects in Scenes

Early studies of context effects on scene perception focused on whether and how a setting influences object recognition. In Palmer's (1975) well-known study, information was available before viewing to provide the context for object recognition. Participants viewed a line drawing of a picture for 3 sec, followed by the presentation of an object for 20, 40, 60, or 120 msec. The participants identified objects most accurately when they were preceded by a related scene and least accurately when they were preceded by an unrelated scene. However, since the objects did not appear in the scene itself, it was unclear whether similar context effects would occur when the object and the scene appeared simultaneously and briefly.

Context effects also have appeared in eye-tracking studies with long viewing times. In eye-tracking studies, the length of fixation has been taken to reflect the amount of processing. Friedman (1979) found that inconsistent objects were fixated longer than consistent objects. De Graef, Christiaens, and d'Ydewalle (1990) found a similar effect, but only after multiple fixations had been made in the scene. However, since fixations typically last on the order of 300 msec and the meaning of a scene can be extracted in about 100 msec (e.g., Intraub, 1981; Potter, 1975; Thorpe, Fize, & Marlot, 1996), it is possible that longer fixation durations reflected additional processing unrelated to object identification. For instance, inconsistent objects may be more interesting or novel.

In another line of work, an object detection task has been used to determine whether objects in plausible settings would be detected more readily than objects in implausible settings. In the object detection task designed by Biederman (Biederman, Mezzanotte, & Rabinowitz, 1982), a name of an object was presented, followed by a brief presentation of a scene. Participants indicated whether the named object had appeared in the scene by responding *yes* or *no*. Studies in which this task has been used have shown that objects in typical settings (e.g., a chicken on a farm) were detected more accurately than were objects in implausible settings (e.g., a chicken in a living room) (Biederman et al., 1982; Boyce & Pollatsek, 1992; Boyce, Pollatsek, & Rayner, 1989). However, once guessing in this paradigm was adequately controlled, scene context did not appear to influence object perception (Hollingworth & Henderson, 1998, 1999).

Hollingworth and Henderson (1998, 1999) failed to find consistency effects in a series of two-alternative forced choice (2AFC) tasks. Participants viewed line drawings of scenes and were asked to select which of two objects had appeared. Since accuracy was similar for objects in consistent and inconsistent scenes, Hollingworth and Henderson proposed the functional isolation model of scene perception. This account suggests that objects in scenes are

J. L. Davenport, jdavenport@cmu.edu

 Copyright 2007 Psychonomic Society, Inc.

processed independently from their backgrounds and that context effects arise from deliberate reasoning processes.

More recently, Davenport and Potter (2004) broadened the investigation of consistency effects, using manipulated color photographs containing a salient foreground object. The pictures were presented for 80 msec, followed by a visual mask. A naming task was used so that no additional information would be available during the trial and no objects or backgrounds would be repeated. The results, that objects were perceived more accurately in consistent settings and that backgrounds were perceived more accurately with consistent foreground objects, led Davenport and Potter to propose an interactive account of object and scene processing. Contrary to a strict version of the functional isolation model, the interactive model suggests that scenes are processed holistically. Objects and their settings are processed together and mutually constrain each other.

Unlike prior models of contextual processing (Biederman et al., 1982; Boyce et al., 1989; Boyce & Pollatsek, 1992; Friedman, 1979; see Henderson & Hollingworth, 1999, for a review), the interactive model makes no assumption that the background context or scene schema is computed first. Instead, contextual information is available from both objects and their settings: Objects provide the context for the background, and the background provides the context for foreground objects.

The interactive model implies that objects, as well as backgrounds, produce contextual effects on perception. However, in Davenport and Potter's (2004) experiments, stimuli contained only one central, foreground object in a background. Since the prior studies of Hollingworth and Henderson (1998, 1999) contained multiple objects, it is possible that the consistency effect found by Davenport and Potter was specific to single-object scenes. Furthermore, with only a single object present, the studies could not test the interactive account's prediction that, in addition to objects and backgrounds' influencing each other, objects may also influence one another.

Related Objects Outside of Scenes

Related objects have been shown to facilitate each other's perception when presented outside of scenes or with long viewing times in scenes. Henderson, Pollatsek, and Rayner (1987) found that priming from a previously fixated object shortened the naming latency for a related object. De Graef et al. (1990) speculated that context effects in a free-viewing task might be due to priming between objects in scenes. More recently, Auckland, Cave, and Donnelly (in press) presented related objects in a circular layout with no background scene. The central target object appeared either concurrently with or following the related objects. In a six-alternative forced choice paradigm, target objects (e.g., a hand of cards) were identified more accurately when surrounded by semantically related objects (e.g., pictures of dice, dominos, and poker chips) than when surrounded by semantically unrelated objects (e.g., different types of fruit). A group of semantically related objects outside of a scene context can influence the perception of a single target object.

Related Objects in Scenes

Since prior studies suggest that objects may influence the perception of their background and that related objects influence object perception in arrays, it seems plausible that objects in scenes may influence each other in a briefly presented scene. However, the single prior study addressing how objects interact in scenes did not show this to be the case. Boyce et al. (1989) manipulated whether a target object in a scene was presented with episodically related or unrelated cohort objects. Contrary to an interactive account of processing, no influence of object-to-object context was found, although the scene context did have an effect. One possible explanation for the lack of object-to-object effects is that the sparse and relatively small line drawings used as stimuli may have made the objects difficult to interpret outside of their scene context.

In the present experiments, we investigated consistency effects between objects in scenes viewed a single time for brief durations. Any influence of context in this design would need to occur in a single glimpse of a scene. Experiment 1 tested whether the presence of related foreground objects would eliminate consistency effects with the background. If the background influenced the foreground object only because the attentional load was low when a single object was present, adding a second object to be reported should increase processing and diminish the influence of the background. If the interactive model is correct, however, perception of an object should be influenced by its setting even when an additional, related object is present. Furthermore, the interactive account suggests that foreground objects will influence background perception even when the objects do not appear at fixation. In Experiment 2, we investigated the effects of related foreground objects on background perception. Finally, the interactive account predicts that objects influence each other's perception. In Experiment 3, the relatedness between foreground objects was manipulated to determine whether relatedness and background consistency would make independent contributions to object perception and whether object-to-object consistency effects would occur with a single related object.

EXPERIMENT 1

Using brief, masked presentations of color photographs of backgrounds with a single foreground object, Davenport and Potter (2004) found that objects were reported more accurately when their settings were semantically consistent than when they were inconsistent. Would the background still affect object perception when an additional, related object was present? In most normal circumstances, scenes contain more than a single salient foreground object, and in such cases, the influence of the background on object perception may be less important than the relations among foreground objects. In prior studies (e.g., Hollingworth & Henderson, 1998), the influence of the background on object perception might have been absent because related objects were present in the scene. Also, if there were two or more objects in the foreground, the added attentional load might have reduced the influence of the background on object perception.

In Experiment 1, participants were instructed to report the prominent foreground object or objects in a briefly presented scene. Scenes had either one or two foreground objects. If two objects were present, they were always episodically related to each other; that is, they were both plausible in the same setting. The background setting was either semantically consistent or inconsistent with the object(s). These manipulations enabled a test of whether the number of foreground objects would modulate the object–background consistency effect. Two foreground objects would occlude more of the background and increase the attentional load, possibly eliminating the contextual influence from the setting. However, if the holistic, interactive account is correct, the effect of object–background consistency should remain when two related foreground objects were present—even when less of the background was visible.

Method

Participants. Sixteen fluent English speakers with normal or corrected-to-normal vision from the Massachusetts Institute of Technology community volunteered and were paid for their participation.

Materials and Apparatus. The stimuli consisted of 40 color photographs of diverse settings. For each background image, two objects that would be likely to appear in that setting were selected from different source photographs. Although each object was consistent with the scene, the objects themselves were not necessarily strong associates (see Table 1 for a list of objects and backgrounds). In this article, objects are termed *related* if they fit into the same consistent scene, whether they were presented to a given participant with that scene or a different, inconsistent scene. The backgrounds and objects were taken from commercially available photo CDs and the Web. Objects were people, animals, furniture, vehicles, and the like and ranged in size from 34 × 64 pixels to 391 × 156 pixels.

All image manipulation was performed using Adobe Photoshop 7.0. To create the consistent stimuli, the two related objects were pasted into the *consistent* background. To create the inconsistent stimuli, the background photos were paired, and the objects were exchanged between scenes. For example, a sofa and a lamp were consistent in an apartment setting but inconsistent in a street setting. The foreground objects appeared in the same locations in each picture and were pasted so that size and support relations were not violated (see Figure 1 for example stimuli). To create the one-object condition, one foreground object was removed without changing the position of the other object. Whether a given setting appeared with one or two objects, which object was presented in the one-object condition, and whether these objects were consistent or inconsistent with the setting were fully counterbalanced between participants. A set of masks was generated by cutting each of six other pictures into a 20 × 20 grid of rectangles and randomly rearranging the pieces.

All the pictures and masks consisted of jpeg files 500 pixels in width × 300 pixels in height. They were presented on an Apple PowerMac G3 computer with a 400-MHz processor. The 17-in. monitor was set to a resolution of 1,024 × 768 pixels, with a refresh rate of 75 Hz. As displayed, the pictures were 17.64 × 10.53 cm, subtending approximately 22º of visual angle horizontally and 13º of visual angle vertically when viewed from a normal viewing distance of 45 cm. The experiments were written in MATLAB, using the Psychophysics Toolbox extensions (Brainard, 1997; Pelli, 1997). The pictures appeared on a black background that was present throughout the experiment, and the room was normally illuminated.

Design and Procedure. Each of the 40 trials consisted of a single picture followed by a mask. Each participant saw half of the pictures in the consistent condition and half in the inconsistent condition. Of these, half the pictures contained a single object, and half contained two objects from the same setting. The pictures were randomly intermixed, and the participants never saw an object or a background more than once.

Each trial consisted of a central fixation "+" for 300 msec, a blank of 200 msec, the test picture for 80 msec, and a mask for 200 msec. A dialog box appeared immediately after the mask. In Experiment 1, the dialog box had a single entry area on one-object trials and two entry areas on two-object trials. The participants did not know in advance whether the picture on a given trial would contain one or two objects.

The participants were informed that each picture would contain one or two objects that might or might not belong with the background. Their task was to type the name of the foreground object(s) into the response box. The participants were to type "?" if they did not see the object(s). At the end of the experiment, the participants were shown each picture again for 500 msec without a mask and

Table 1
Backgrounds and Objects

Background	Objects	Background	Objects
Farm	pig, tractor	Arctic	Eskimo, igloo*
Classroom	teacher, overhead projector*	Yard	lawnmower, wheelbarrow
Warehouse	crates, forklift*	Forest	bear, moose
Intersection	ambulance, traffic cone	Desert	rock, cactus
Bed	cat, teddy bear	Beach	sandcastle, beach ball*
Park	stroller, park bench	Stage	piano, cello*
Pool	lifeguard, pool chair	Track	runner, hurdle*
Hospital	doctor, nurse*	Church	priest, nun*
Sky	biplane, helicopter*	Outer space	astronaut, satellite
Fireplace	broom, logs	Undersea	turtle, fish
Ice rink	hockey player, goal*	Ocean	buoy, sailboat
Snowy mountain	sledder, skier	Lake	jet ski, motorboat*
Pond	duck, frog	Stove	kettle, frying pan*
Serengeti	zebra, photographer	Library	armchair, book cart
Parking lot	motorcycle, car*	Arena	matador, bull*
Apartment	sofa, lamp	Trail	cyclist, jogger
Patio	grill, patio table	Soccer field	soccer player, soccer ball*
Football field	football player, referee*	Hotel lobby	bellhop, luggage*
Restaurant	chef, waiter*	Basketball court	hoop, basketball player*
Tree	squirrel, bird	Bar	wine, wineglass*

Note—Items on the same line were paired; items with asterisks were rated as strong associates in a post hoc analysis.

Figure 1. Example stimuli. Top, consistent; middle, inconsistent; bottom (Experiment 3 only), unrelated. In Experiments 1 and 2, one object was removed on half of the trials.

then wrote the name of the object or objects and the background setting.

Scoring. All the results were scored blind to condition. The responses were scored as correct if they were the same names as those provided by the participants in the postexperimental naming session or synonyms at an equal level of descriptiveness (e.g., *runner* and *jogger*). Since the participants gave specific names for over 98% of the objects and settings in the postexperimental trials, an unmasked presentation of 500 msec appeared to be adequate for full processing of these stimuli. The responses were scored as incorrect if they were names different from or more general than (e.g., *animal* instead of *pig*) those given in the naming session or if the participants responded with a question mark.

If a participant guessed the object that would have been consistent when the inconsistent scene was presented (e.g., responded *igloo* when a pig was presented in the arctic scene), the response was considered an intrusion. To correct for such pure guesses based on the background, for each intrusion made by a given participant, one correct consistent response was subtracted. Such intrusions were extremely rare. Intrusions occurred in 0.6% of the responses in Experiment 1. All analyses were carried out on the corrected data (see Table 2 for a complete breakdown of error responses for all the experiments). Note that consistent with Grill-Spector and Kanwisher (2005), viewers rarely reported the general category of an object rather than its specific identity.

Results and Discussion

An ANOVA was carried out to determine whether accuracy in reporting an object varied as a function of background consistency (consistent vs. inconsistent) and number of objects (one vs. two). A highly significant main effect of object–background consistency was found, with objects in consistent settings ($M = .74$) reported more accurately than objects in inconsistent settings ($M = .59$) [$F(1,15) = 49.96$, $p < .001$]. No main effect of number of objects was found ($p = .13$), and the interaction of consistency and number of objects did not approach significance [$F(1,15) < 1$]. The results are shown in Figure 2. The same pattern of effects was found in an analysis in which items was used as a random variable. A significant main effect was found for consistency [$F(1,79) = 14.71$, $p < .001$]. No significant difference was found for either the main effect of number of objects or the interaction.

Consistent with the interactive account of scene processing, objects were reported more accurately in consistent than in inconsistent settings. Although the background could have been ignored in Experiment 1, the presence of an additional, related object did not diminish the influence of the background on object perception. Accuracy in reporting two objects was not significantly different from the accuracy in reporting a single object, despite the very brief presentation duration and performance that was well below ceiling (in a post hoc analysis, this experiment had a power of .33 to detect an effect). Whether the relatedness of the two objects contributed to the ease of reporting them was a question addressed in Experiment 3.

EXPERIMENT 2

Experiment 1 demonstrated that consistency effects on object perception were not modulated by the presence of an additional foreground object. Experiment 2 tested whether consistency effects on background perception would be affected by the number of foreground objects.

Table 2
Number and Proportion of Each Type of Error for Each Condition

Scene	Object Condition	? Responses No.	? Responses Prop.	Specific Incorrect No.	Specific Incorrect Prop.	Vague No.	Vague Prop.	Correct Generic No.	Correct Generic Prop.	Intrusions No.	Intrusions Prop.	Total Errors
Experiment 1: Report Object or Objects												
Consistent	One object	9	.24	23	.62	–	–	4	.11	–		37
	Two objects	36	.38	52	.55	–	–	5	.05	–		95
Inconsistent	One object	10	.15	45	.70	1	.02	7	.11	1	.02	64
	Two objects	57	.43	57	.41	1	–	19	.14	2	.01	138
Experiment 2: Report Background												
Consistent	One object	8	.22	18	.50	2	.06	8	.22	–		36
	Two objects	9	.24	22	.58	–	–	7	.12	–		38
Inconsistent	One object	20	.27	41	.55	–	–	9	.12	5	.07	75
	Two objects	19	.22	43	.51	1	.01	15	.18	7	.08	85
Experiment 3: Report Objects												
Consistent	Related	44	.44	48	.48	1	.01	8	.08	–		101
	Unrelated	50	.42	59	.50	1	.01	8	.07	–		118
Inconsistent	Related	58	.43	64	.47	4	.03	8	.06	1	.01	135
	Unrelated	61	.41	72	.49	1	–	13	.09	1	.01	148

Note—? Responses, those with "?" or left blank; Specific Incorrect, incorrect responses at a specific level of detail (e.g., *kitchen*); Vague, incorrect responses without detail (e.g., *weird thing*, *blue*, *something*); Correct Generic, incorrect responses due to insufficient detail (e.g., *animal*, *outside*); Intrusions, incorrect responses that matched correct consistent responses. When two objects were to be reported, the opportunity for errors was doubled.

If background perception is influenced by the amount of space taken up by foreground objects, participants should be less accurate in reporting the background when two objects, rather than one, appear in the foreground. A further question was whether the object–background consistency effect would be magnified in the two-object condition. In Experiment 3 of Davenport and Potter (2004), backgrounds with a single consistent foreground object were identified as accurately as backgrounds with no foreground objects. However, two related objects consistent with the background might make accurate perception of the setting more likely, whereas two inconsistent objects would have the opposite effect.

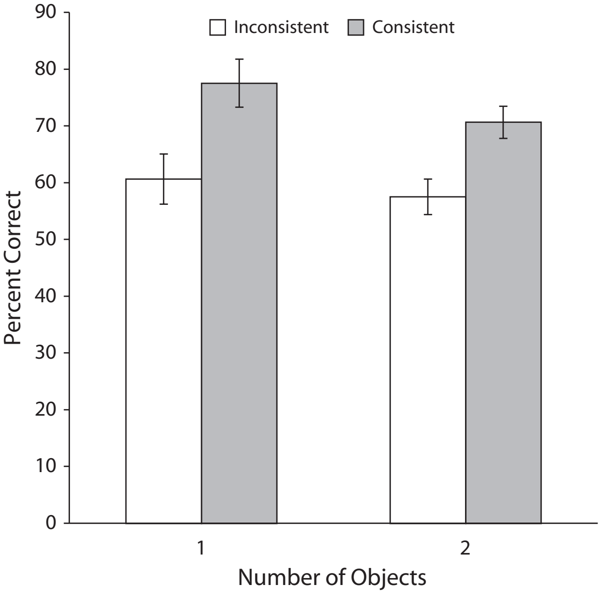

Figure 2. Experiment 1 results: Accuracy in reporting objects.

In Experiment 2, the participants reported the background setting of each scene, which contained either a single foreground object or two related foreground objects that were either consistent or inconsistent with their setting.

Method

The method in Experiment 2 was identical to that in Experiment 1, except as noted. Sixteen new subjects participated in this experiment.

Design and Procedure. In Experiment 2, the participants were instructed to name the background setting or type of place for each picture. They were informed that each picture would contain one or two objects that might or might not belong with the setting but that their task was to type the name of the background into the response box or to type "?" if they did not see the background.

Scoring. If a participant guessed the background that would have been consistent when the inconsistent scene was presented (e.g., responded *farm* when a pig was presented in the arctic scene), the response was considered an intrusion. To correct for such pure guesses based on the foreground objects, for each intrusion made by a given participant, one correct consistent response was subtracted. All analyses were carried out on the corrected data. Intrusions occurred in 4% of the responses in Experiment 2.

Results and Discussion

An ANOVA was carried out to determine whether accuracy in reporting the background would vary as a function of consistency or the number of foreground objects. A strong main effect of object–background consistency was found, with backgrounds reported more accurately when they contained consistent foreground object(s) ($M = .72$) than when they contained inconsistent foreground object(s) ($M = .50$) [$F(1,15) = 49.09, p < .001$]. Neither the main effect of number of foreground objects ($F < 1$)

nor the interaction of consistency and number of objects ($F < 1$) approached significance. The results are shown in Figure 3. The same pattern of effects was found in an analysis carried out with items as a random variable. Although the consistency effect was significant [$F(1,39) = 27.93$, $p < .001$], there was no significant main effect of number of objects and no significant interaction.

In Experiment 2, background report was influenced by the consistency of the foreground object(s), but not by the number of foreground objects. Object–background effects were robust and were not modulated by the presence of an additional object. Two related objects had the same influence on background perception as a single object, and the sheer amount of space taken up by two foreground objects did not appear to influence perception.

EXPERIMENT 3

Do objects influence one another's perception? If object perception and scene perception are truly interactive, not only should objects and backgrounds influence each other, but also objects should influence other objects. If object-to-object consistency plays a role in perception, an object should be identified more accurately in the presence of a related object than in the presence of an unrelated object. Experiment 1 could not speak to whether object relatedness influences perception, since the two objects were always related. Furthermore, in the two-object condition, the benefits of relatedness were pitted against the requirement of reporting two versus a single object on those trials. In Experiment 3, the participants saw and attempted to name two foreground objects on each trial. On half of the trials, the objects were related, so that both or neither was consistent with the background. On the other half of the trials, the objects were unrelated; one object was consistent with the background, the other object was inconsistent.

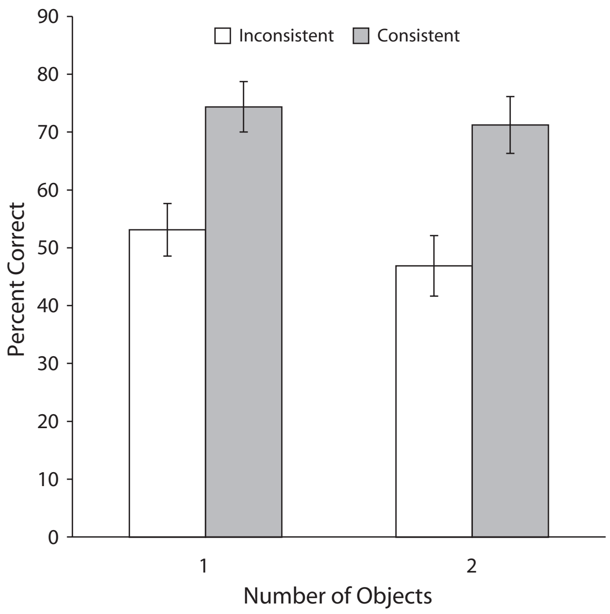

Figure 3. Experiment 2 results: Accuracy in reporting background settings.

Method

The method of Experiment 3 was identical to that in Experiment 1, except as noted. Sixteen participants who had not taken part in either of the previous experiments participated.

Design and Procedure. The same backgrounds and objects as those in Experiments 1 and 2 were used in Experiment 3. Two objects appeared in each picture. On related trials, the objects were both consistent or both inconsistent with the background, as in Experiment 1. On unrelated trials, one object was consistent with the background, and one was inconsistent. For example, the classroom and yard scenes were paired to create unrelated trials in which the classroom scene contained a wheelbarrow and an overhead projector and the yard scene contained a lawnmower and a teacher. Whether a given setting appeared with two consistent objects, two inconsistent objects, or one consistent and one inconsistent object was fully counterbalanced between participants. The participants were instructed to report the two objects on each trial.

Scoring. Intrusions occurred in 0.4% of the responses in Experiment 3.

Results and Discussion

An ANOVA was carried out on responses to objects, separately for each object, with object–background consistency and object relatedness as variables. There were significant main effects of object–background consistency and object relatedness. Objects in a consistent setting ($M = .67$) were reported more accurately than were objects in an inconsistent setting ($M = .57$) [$F(1,15) = 11.73$, $p < .01$]. Related objects were reported more accurately ($M = .66$) than were unrelated objects ($M = .59$) [$F(1,15) = 5.57$, $p < .05$]. Object–background consistency and object relatedness did not interact ($F < 1$). The results are shown in Figure 4. A similar pattern of effects was found in an ANOVA carried out with items as a random variable. Significant main effects were found both for consistency [$F(1,79) = 9.09$, $p < .01$] and for relatedness [$F(1,79) = 9.35$, $p < .01$]. No significant interaction was found.

Consistent with an interactive account of scene processing, the results suggest that object perception is influenced both by the background setting and by the presence of a related object. Note that object relatedness is defined in this experiment as potential co-appearance in a given setting; related objects were not selected with the criterion of being strong associates.

To determine whether the strength of association between related items influenced the relatedness effect, 5 new participants were given the names of the related pairs of items and used a Likert scale to rate the strength of association from 1 (*not at all associated*) to 7 (*strongly associated*). The mean rating was 4.78. The 20 pairs with the highest means ($M = 5.92$) were coded as *strong* associates, and the remaining 20 pairs ($M = 3.64$) were coded as *weak* associates (the 20 strong associates have asterisks in Table 1). An ANOVA was carried out with consistency, relatedness, and strength of association (strong vs. weak) as variables. In addition to the previously mentioned main effects of consistency and relatedness, there was a signifi-

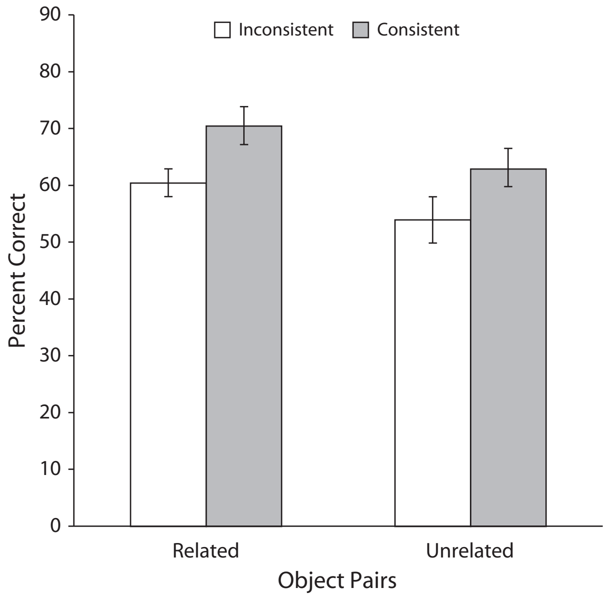

Figure 4. Experiment 3 results: Accuracy in reporting objects.

cant interaction of relatedness with strength of association [$F(1,15) = 5.68$, $p < .05$] but no significant main effect of association strength or any other interaction. Planned comparisons were carried out separately for strong and weak association strengths. A significant main effect of relatedness was found for strong associates, with related items reported more accurately ($M = .71$) than unrelated items ($M = .58$) [$F(1,15) = 8.97$, $p < .01$]. However, no significant main effect of relatedness was found for weak associates. Related objects ($M = .62$) were not reported significantly more accurately than unrelated objects ($M = .60$) ($F < 1$). Further research is needed to determine the relative contributions of semantic association and episodic relatedness to these object-to-object effects.

GENERAL DISCUSSION

Three experiments tested the role of semantic consistency in object and background perception. In Experiment 1, objects in consistent scenes were reported more accurately than were objects in inconsistent scenes, regardless of the number of foreground objects. In Experiment 2, backgrounds with consistent foreground objects were reported more accurately, regardless of the number of foreground objects. In Experiment 3, in addition to a consistency effect, objects appearing with a related object were reported more accurately than were objects appearing with an unrelated object, regardless of background. The two effects did not interact. Together, the findings support a model of scene processing in which objects and scenes are processed interactively, with context effects for both object-to-object relatedness and object–background consistency.

The strong object–background consistency effects are in line with prior findings (Biederman et al., 1982; Boyce & Pollatsek, 1992; Boyce et al., 1989; Davenport & Potter, 2004). Attention to multiple foreground objects did not eliminate the influence of the background, even when the background could have been ignored. A post hoc analysis of object size in Experiments 1 and 3 showed that although larger objects were reported more accurately than were smaller objects ($p < .001$), the consistency effect was similar across objects. Thus, the object–background consistency effects are robust: The ability to report objects is influenced by their settings.

Experiment 3 demonstrates for the first time that related objects in scenes influence one another's perception. These results extend a prior finding of semantic relatedness between objects in arrays (Auckland et al., in press) to objects in scenes. In Auckland et al.'s experiments, multiple related objects that were strong semantic associates of each other were used (e.g., *hand and foot* or *hammer, screwdriver, nails, and pliers*), as in Henderson et al.'s (1987) study in their object–object priming experiment. In the present experiments, related objects were not deliberately chosen to be strong associates of each other. Although an item analysis revealed a significant main effect of relatedness across items, a post hoc analysis suggested that object-to-object effects were greater for strongly associated objects than for weakly associated objects.

The object-to-object context effects in Experiment 3 conflict with the findings of Boyce et al. (1989), who failed to find an effect of relatedness between objects in scenes. However, as has been noted, the stimuli in that study were sparse line drawings; the objects may have been difficult to interpret. Also, because the task was to detect the presence of an object named in advance, the task might not have been sufficiently sensitive to detect effects of object relatedness. Furthermore, Boyce et al. defined related objects as those likely to appear in the same scene. These episodically related objects may not have been as strongly associated as the related objects used in the present study.

The results are inconsistent with a strict version of the functional isolation model of scene processing proposed by Hollingworth and Henderson (1998, 1999) after they failed to find consistency effects in several experiments. The functional isolation model proposed that context does not influence the processing of objects in scenes. The discrepancies between the results of the present experiments and the results in Hollingworth and Henderson may have been due to the task, the design, or the stimuli. In their experiments, the task was either object detection, in which participants indicated whether a named object was present in a scene by responding *yes* or *no*, or a 2AFC task following presentation of the scene, in which participants indicated which of two named or pictured objects had been presented. The stimuli in their experiments were black-and-white line drawings, and objects and backgrounds were repeated multiple times.

The present experiments tested whether the semantic consistency between objects and scenes affects conscious perception the very first time a naturalistic picture is seen. The naming task provided an indication of what information the participant was able to extract in a brief, masked presentation of 80 msec with no further information being given on that trial. In contrast, both the object detection and the 2AFC tasks used by Hollingworth and Henderson (1998, 1999) provided information on the trial other than the picture (e.g., the target label in the object detection task or the names in the 2AFC task). Although 2AFC task eliminates concerns about a bias to make a background-consistent response (in that the objects between which the participant made a choice were both consistent or both inconsistent), the additional information on the trial may still have enabled postperceptual reasoning. For example, although the participant may not have extracted enough information to name an object (e.g., a chicken) when given two choices (e.g., a chicken or a pig), the participant may have reflected back on the picture and determined that one choice was more likely on the basis of a partially processed image. The 2AFC task cannot distinguish whether the viewer saw some sort of bird-like object, and so chose *chicken* over *pig*, or actually saw a chicken. In the present experiments, to be counted as correct, the participant had to provide the name of the object at the basic level (e.g., *Eskimo*, not *person*, and *pig* rather than *animal*), without being aided by having alternatives to choose between. Since neither the 2AFC nor the present naming paradigm can determine when in processing consistency effects occur (or fail to occur), future studies are needed to determine the fine-grained time course of contextual processing.

Furthermore, Hollingworth and Henderson (1998, 1999) may have failed to find consistency effects because the same objects appeared in different scenes and, if an inconsistent object was present (e.g., a blender in a farm scene), it was always the item later tested. Target objects on consistent trials had no such distinction (e.g., a chicken, tractor, barn, etc. all would appear in a farm scene). Finally, the present experiments used natural color photographs with salient foreground objects, which may convey consistency information more readily than black-and-white line drawings with small objects (Cheng & Simons, 2001).

Are the present results due to response bias? Since no information other than the picture was presented during a trial, the naming task used here was not susceptible to the type of response bias present in earlier object detection studies (e.g., Biederman et al., 1982; Boyce et al., 1989; as pointed out in Hollingworth & Henderson, 1998, 1999). In those studies, the task was to respond *yes* or *no* when a target label (e.g., *sofa*) matched or did not match, respectively, an object in a briefly presented scene (e.g., a sofa in a living room). In Biederman et al.'s (1982) study and the replication by Hollingworth and Henderson (1998, Experiment 1), false alarm rates were higher when the target label (e.g., *sofa*) and pictured scene were consistent (e.g., a living room) than when they were inconsistent (e.g., a street); the use of the combined false alarm rate in calculating d' led to an overestimation of sensitivity in consistent conditions and an underestimation of sensitivity in inconsistent conditions. When Hollingworth and Henderson (1998, 1999) used separate false alarm rates for calculating d' in consistent and inconsistent conditions, the effect of consistency disappeared. However, this type of response bias—a tendency to respond *yes* when the background was consistent with a specified target—was not possible in the present experiments, because no advance target was specified.

Another possible form of bias, however, could arise if the participants processed only the background (in Experiments 1 and 3) or processed only the object(s) (in Experiment 2) and guessed which objects or backgrounds, respectively, were presented. This type of pure guessing was conservatively corrected for before the data analysis. For example, if the task was to report the objects and the participants saw only the background and simply guessed a related object, their responses would have occasionally included intrusions of consistent objects when an inconsistent object was presented (e.g., reporting *lawnmower* when an overhead projector was presented in a yard). Intrusions of this type were very rare, and their potential contribution to correct responses was corrected for by subtracting one correct consistent response for each such intrusion on inconsistent trials. Thus, if the consistency advantage had been the result of a pure guessing strategy, subtracting intrusions from correct consistent responses would have eliminated the consistency advantage.

So how might the consistency effect occur? The present experiments addressed how context influences conscious perception of briefly presented pictures. It seems likely that rather than penetrating and influencing early vision, conceptual knowledge about the world affects later stages of perceptual processing, biasing perceptual interpretation in a Bayesian fashion (e.g., Knill & Pouget, 2004). When viewing time is very short, as in the present experiments, there will be some degree of uncertainty about the background, the objects, or both. The prior likelihood that an igloo is on a farm is low, so more evidence is required before the perceptual system concludes that there actually is an igloo. For a pig, however, the prior likelihood is much higher, so less evidence is required. This kind of bias has been shown to be common in vision—for example, in identifying letters in words (e.g., Massaro & Cohen, 1991). In everyday life, this bias will increase accuracy in most situations, allowing us to recognize objects with less information than would be needed to recognize an unexpected object—at the cost of making occasional errors when the object is, in fact, unlikely in that setting.

Bar (2004) has proposed a model of contextual facilitation that speculates about how individual object perception may be influenced by context. In this model, two types of information, a context frame and candidate objects, are processed in parallel and used to constrain the interpretation of an object. The context frame, on the basis of global, low-spatial-frequency information, provides prototypical information about what objects usually occur and where. The candidate objects, on the basis of local information, provide shape-based options for possible objects. Bar proposed that the context frame is generated first in the parahippocampal cortex, whereas candidate objects are generated more slowly in the prefrontal cortex. Furthermore, these two sources of information are combined in the inferior temporal cortex, and their intersection produces reliable selection of object identity. The interactive account proposed by Davenport and Potter (2004) is compatible with Bar's contextual facilitation model but would not require the additional assumption that backgrounds need to be processed prior to objects.

Can this inherent perceptual bias toward the more probable interpretation—what Helmholtz (1925) called "unconscious inference"—be distinguished from another, more conscious kind of bias often called "sophisticated guessing"? For example, in the latter case, the viewer sees a farm scene with something that looks like an animal, reasons that it is more likely to be a pig than a wolf or a hyena, and so simply guesses that it is a pig. Both the sophisticated-guessing hypothesis and the Bayesian model assume a bias toward the probable, but the Bayesian claim made here is that the bias is built into perception, rather than being a conscious guessing strategy subsequent to perception. Because the participants in the present experiments were asked to report what they saw and to enter a question mark if they did not know, and because they rarely came up with a guess (in the inconsistent condition) that matched the corresponding item in the consistent condition, it seems likely that the effects of consistency and relatedness were unconscious, affecting what the participants perceived, rather than what they simply guessed. However, since the present paradigm cannot distinguish whether participants were using the same threshold for responding in the consistent and the inconsistent conditions, further work will be required to show conclusively that these effects occur before conscious perception.

The present results suggest an interactive model of object perception in scenes in which both object–background and object–object effects occur. The gist or meaning of a scene may be extracted from a brief glimpse of 100 msec (Intraub, 1981; Potter, 1975; Thorpe et al., 1996), leading prior models of semantic consistency in scene perception to focus on how background information influences object perception (Biederman et al., 1982; Boyce & Pollatsek, 1992; Friedman, 1979). The present study has also looked at how an object influences the perception of another object, in the presence of a consistent or inconsistent background. An interactive model of scene processing suggests that scenes are processed holistically, with mutually constraining object and background processing. Scenes containing related objects and consistent settings may require less perceptual information for identification of the individual elements. In summary, the present study shows that backgrounds influence how objects are perceived and that objects influence the perception of other objects and their backgrounds.

AUTHOR NOTE

This research was supported by Grant MH47432 from the National Institute of Mental Health. I thank Mary C. Potter for helpful discussions throughout the course of this work and Darlene Ferranti and Laura Fox for assistance with stimulus creation and programming. Correspondence concerning this article should be addressed to J. L. Davenport, Department of Psychology, Carnegie Mellon University, Baker Hall 354J, 5000 Forbes Avenue, Pittsburgh, PA 15213 (e-mail: jdavenport@cmu.edu).

REFERENCES

AUCKLAND, M. E., CAVE, K. R., & DONNELLY, N. (in press). Nontarget objects can influence perceptual processes during object recognition. *Psychonomic Bulletin & Review*.

BAR, M. (2004). Visual objects in context. *Nature Reviews Neuroscience*, **5**, 617-629.

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, & 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 & Performance*, **15**, 556-566.

BRAINARD, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, **10**, 433-436.

CHENG, E. K., & SIMONS, D. J. (2001, November). *Perceiving the internal consistency of scenes*. Paper presented at the 42nd Annual Meeting of the Psychonomic Society, Orlando, FL.

DAVENPORT, J. L., & POTTER, M. C. (2004). Scene consistency in object and background perception. *Psychological Science*, **15**, 559-564.

DE GRAEF, P., CHRISTIAENS, D., & D'YDEWALLE, G. (1990). Perceptual effects of scene context on object identification. *Psychological Research*, **52**, 317-329.

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

GRILL-SPECTOR, K., & KANWISHER, N. (2005). Visual recognition: As soon as you know it is there, you know what it is. *Psychological Science*, **16**, 152-160.

HELMHOLTZ, H. (1925). *Physiological optics: Vol. III. The perceptions of vision* (J. P. C. Southall, Trans.). Rochester, NY: Optical Society of America.

HENDERSON, J. M., & HOLLINGWORTH, A. (1999). High-level scene perception. *Annual Review of Psychology*, **50**, 243-271.

HENDERSON, J. M., POLLATSEK, A., & RAYNER, K. (1987). Effects of foveal priming and extrafoveal preview on object identification. *Journal of Experimental Psychology: Human Perception & Performance*, **13**, 449-463.

HOLLINGWORTH, A., & HENDERSON, J. M. (1998). Does consistent scene context facilitate object perception? *Journal of Experimental Psychology: General*, **127**, 398-415.

HOLLINGWORTH, A., & HENDERSON, J. M. (1999). Object identification is isolated from scene semantic constraint: Evidence from object type and token discrimination. *Acta Psychologica*, **102**, 319-343.

INTRAUB, H. (1981). Rapid conceptual identification of sequentially presented pictures. *Journal of Experimental Psychology: Human Perception & Performance*, **7**, 604-610.

KNILL, D. C., & POUGET, A. (2004). The Bayesian brain: The role of uncertainty in neural coding and computation. *Trends in Neurosciences*, **27**, 712-719.

MASSARO, D. W., & COHEN, M. M. (1991). Integration versus interactive activation: The joint influence of stimulus and context in perception. *Cognitive Psychology*, **23**, 558-614.

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

PELLI, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, **10**, 437-442.

POTTER, M. C. (1975). Meaning in visual search. *Science*, **187**, 965-966.

THORPE, S., FIZE, D., & MARLOT, C. (1996). Speed of processing in the human visual system. *Nature*, **381**, 520-522.

(Manuscript received September 27, 2005;
revision accepted for publication January 6, 2006.)