This is to certify that the

dissertation entitled

The Effects of Scene Context on Perceptual Encoding: Evidence From a New Paradigm

presented by

Phillip Anthony Weeks, Jr.

has been accepted towards fulfillment of the requirements for

Ph.D. degree in Psychology

Major professor

Date September 12, 1995

THE EFFECTS OF SCENE CONTEXT ON PERCEPTUAL ENCODING: EVIDENCE FROM A NEW PARADIGM

Phillip Anthony Weeks, Jr.

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Department of Psychology

1995

ABSTRACT

THE EFFECTS OF SCENE CONTEXT OF PERCEPTUAL ENCODING:
EVIDENCE FROM A NEW PARADIGM
By

Phillip Anthony Weeks Jr.

Previous research examining the effects of scene context on perceptual encoding has been criticized for possibly reflecting post-perceptual encoding. In the research reported here, a same/different decision task was used to examine these effects in an attempt to circumvent some of the criticisms of earlier experiments. Participants were shown a study scene followed by a mask and then a test scene, and determined if a target object had undergone either a deletion or an orientation change. The study scene was presented for 250, 500, or 2500 msec. The results of Experiments 2,3, and 4 showed effects of scene context on detection of target object manipulations. An eye-movement monitoring study also found effects of scene context on various eye-movement measures. These findings are discussed in terms of the perceptual schema hypothesis that states that the gist of a scene guides subsequent perceptual encoding.

ACKNOWLEDGMENTS

I would like to give very special thanks to Gary Schrock, Steve Pierce, and Karen Butler for their invaluable contributions to data analyses, stimuli construction, and participant running on this project, and without whom, this work would have been a long time coming.

This research was partially supported by Grant DAAH04-94-G-0404 from the Army Research Office, Department of the Army, awarded to John M. Henderson.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

As you look around, reflect on of how effortlessly you can make sense of your visual environment. What is quickly apparent is how proficient the visual system is at synthesizing the myriad of visual stimuli that lands on the retina. Psychologists have classified two types of processes performed by the visual system: data-driven processes and conceptually-driven processes. Data-driven processes, or bottom-up processes as they are sometimes called, refer to the processes that begin with the registration of sensory information on the retina and proceed up the visual pathways to higher cortical areas in the brain. Conceptually-driven, or top-down processes, are those processes that use knowledge such as past experiences, expectations, or knowledge about the surrounding context or situation to guide an active search for certain patterns in the visual input (Cowen, Porac, and Ward, 1984). Data-driven and conceptually-driven processes work together or against one another to produce the stable percept we experience when we open our eyes and process the visual environment. This stable percept comprises the meaning we extract about the scene and the objects that we are able to identify in the scene.

As stated above, context, or information about the meaning of a visual scene, also called the gist of a scene, is one example of the type of information used during top-down visual processing. As a result, the effect of context on perception has been an important question for years. It is clear that one's ability to process an object is influenced by the object's context (Sekuler and Blake, 1990). According to Sekuler and Blake, contextual information is more important to visual processing under conditions where the visual input is degraded in some manner. But even under ideal conditions, contextual information is influential during object processing. This latter situation is the focus of the present research.

What type of information the visual system extracts from a scene on a given fixation is heavily researched and debated. Researchers have concluded that some types of information are extracted from scenes very quickly. For example, low-level information such as contrasts in brightness and changes in contour are extracted during initial fixations (Biederman, 1987; Biederman and Ju, 1988). Moreover, a great deal of research suggests that the meaning or "gist" of the scene is also rapidly extracted from a scene (Antes, 1974; Biederman, 1987; Freidman, 1979; Boyce and Pollatsek, 1991). Given that this type of higher level information is quickly extracted, the question then becomes, is this information used to guide object perceptual processing? Numerous studies have shown that scene context information can influence object processing, especially in cases of degraded stimulus information (Sekuler and Blake, 1990). However, the effects of scene context information on object processing under normal, non-degraded stimulus input is not fully understood. Complete object processing combines both perceptual encoding processes and post-perceptual encoding processes. Perceptual encoding can be viewed as the visual processing that takes place from the initial encoding of features of an object in

Post-perceptual encoding processing, then, is all visual processing that takes place after this information has been matched against some memory representation. Thus, the two important issues here are 1.) How quickly can information about the meaning of a scene be apprehended from a scene, and 2.) If this information is apprehended quickly, can it influence subsequent perceptual encoding.

Various studies have shown that scene context has some type of influence on object processing. However, there is still a lack of a consensus on the nature of these effects. One widely held belief is that scene context guides perceptual encoding, or those processes that take place up until the visual stimulus has been matched against its stored memory representation. According to the perceptual schema hypothesis (Biederman, 1981; Loftus and Mackworth, 1978, Friedman, 1979; Boyce, Pollatsek, and Rayner, 1989; Boyce and Pollatsek, 1992), the perception and identification of an object are facilitated by congruent scene context. Other research, however, has shown that scene context does not guide perceptual encoding but influences later object processing (De Graef, Christiaens, and d'Ydewalle, 1990). Unfortunately, disagreements about the effects of context appear to be influenced by the type of experimental paradigm used.

In experiments conducted to determine the effects of frames on object processing, Friedman (1979) found that frames guided object encoding and memory for pictures of real scenes. Frame theories in general describe the representation and use of knowledge for pattern recognition. Frames are abstract representations of knowledge about the world that are obtained through experience, structured in different levels, and are invariant over time. They contain information about the category of a scene and the types of objects that are expected to appear in a given scene. As such, frames represent different types of knowledge stored in an abstract format, a format different from the sensory or linguistic information used to acquire them (Friedman, 1979). According to Friedman, frames, once they are evoked, are used as semantic pattern detectors and can guide subsequent visual processing. This influence is viewed as facilitory if the frame and the object are congruent, and inhibitory otherwise. Thus, an object that is in a strange or inappropriate context will require more processing to identify. In this case, processing must rely on specific features (e.g., lines, corners, etc.) of the object instead of more global features like what type of environment the object appears in. In terms of memory for scenes, Friedman posits that frames serve as a type of heuristic by which information about some earlier presentation of a scene is "remembered by prototyping." Remembering by prototyping, Friedman states, refers to a type of storage heuristic where no particular note of episodic or descriptive information is made for objects that have a reasonably high a priori probability of being found in a particular place within a frame. In other words, objects that are expected or obligatory in a scene activate the frame for that scene and are remembered as prototypes. In this case, their details are not "encoded," but the object is encoded because it is expected to be in the scene. Nonobligatory objects, on the other hand, do not activate a frame for the scene, and thus require more processing to identify. Importantly, however, because nonobligatory objects do not fit the frame, they tend to tag a particular instantiation of an episode, rendering that instantiation or episode more memorable. Thus, memory for a scene in which an unusual object appears will more than likely be greater.

In an experiment conducted to test the predictions of the frame theory, Friedman (1979) gave participants the name of a particular place (for example, kitchen) and then had them view complex pictures of scenes while their eye-movements were monitored. In the scenes, obligatory and nonobligatory objects were manipulated and it was predicted that fixation duration and memory for detail about the object would vary as a function of whether or not the object was an obligatory or nonobligatory object in the scene. Specifically, obligatory objects would be identified faster than nonobligatory objects, figurative detail of nonobligatory objects would be remembered more than the figurative detail of obligatory objects, and finally, changes to scenes would be noticed better if the changes involved nonobligatory objects. Mean durations of the first and second fixations and the third through *n*th fixations were correlated with the rated probability of the objects occurring in the scene.

Friedman found that most of the variance during the first fixation in the amount of time needed to encode an object was accounted for by its rated probability in the scene, with rated probability accounting for less of the variance with subsequent fixations. In other words, as the rated likelihood of an object appearing in a scene increased, the duration of the first fixation on that object decreased. Objects with a lower likelihood of appearing in the scene had longer first fixation durations. She concluded that having some general knowledge about the context of a subsequent scene allows for the instantiation of a frame which is then used to detect obligatory objects, resulting in shorter fixation durations on those objects. Also, identification of lower probability objects requires more processing (longer first and second fixation durations), but results in greater encoding of figurative details of these objects. Concerning recognition memory for changes to the scenes, Friedman found greater accuracy to detecting manipulations when they involved nonobligatory objects, as predicted by the frame theory. Friedman interpreted these findings as support for a frame theory of object processing and scene encoding, where scene context guides perceptual and memory encoding of scenes.

Other evidence for scene context influencing perceptual encoding comes from a study conducted by Antes (1974) examining participants' eye fixation patterns while they looked at scenes. Previous work had shown that eye-movement patterns during scene viewing are influenced by the amount of information that is conveyed in various locations within the scene (Antes, 1974; Mackworth and Morandi, 1967). In a follow-up study, Antes showed that upon initial presentation of a scene, participants quickly fixate areas of high informativeness followed by a greater proportion of subsequent fixations to less informative areas. In his study, the informativeness of a given area within a picture was determined by how much meaning, in and of itself, a particular unit of the picture conveyed as determined by the experimenter. Unit size was determined by fixation densities from eye-movement records and subtended no less than one degree and no greater than five degrees of visual angle in any direction. In this experiment, participants' eye movements were recorded and fixation location and duration measured as they fixated different areas of the pictures. Location of fixation and fixation duration were then evaluated in accordance with the informativeness rating scale for different locations on the pictures as indicated by a separate group of participants. Antes found that the location of participants' fixations increased in informativeness after the first fixation and peaked at the second fixation. Subsequent fixations were to areas of lesser and lesser informativeness. Moreover, participants tended to make larger eye-movements earlier in picture viewing, followed by smaller subsequent eye-movements. Antes concluded that areas of higher informativeness are quickly fixated and that these areas guide subsequent fixations. What is interesting about these findings is that what is informative about a given scene is closely related to the context or the meaning of the scene. Thus, as informative areas within the scene are quickly fixated, information leading to the formulation of the meaning of the scene is quickly extracted by the viewer. Moreover, this scene context information possibly guides subsequent eye fixations. As such, these findings can be taken as support for the early extraction of the meaning or "gist" of the scene, and possibly as support for an early influence of scene context information on perceptual encoding.

Work by Loftus and Mackworth (1978), examining where observers look during picture viewing, has shown some interesting results regarding how quickly information about the meaning of a scene is apprehended. In this study, they defined informativeness as "the extent to which an object has a low a prior probability of being in a picture given the rest of the picture and the viewer's past history" (p. 566). For example, in a farm scene, a tractor would be a noninformative object, while an octopus, in the same location as the tractor, would be an informative object. Thus, they were interested in seeing whether areas of high informativeness would be fixated earlier and more often than corresponding noninformative areas. Participants were shown pictures of scenes that contained either an informative target object or a noninformative target object while their eye-movements were monitored and ordinal fixation number and fixation duration were recorded.

Loftus and Mackworth (1978) found that informative objects were fixated earlier than noninformative objects. The cumulative probability of fixating informative objects as a function of ordinal fixation number was significantly higher than that for noninformative objects. Additionally, they found that the probability of fixating an informative object on any given fixation was greater than that for a noninformative object. Moreover, they found that fixation durations tended to be longer on informative objects than noninformative objects, and this difference in duration increased with subsequent fixations. Loftus and Mackworth concluded that during the early stages of scene processing, several processes must be occurring: 1.) The rapid determination of the gist of the scene, 2.) At least some partial pattern recognition of objects in the periphery, and 3.) Computation of conditional probabilities that these peripheral objects belong in the scene, given the gist of the scene. Further, there must exist a rapid peripheral processing based on cognitive information which determines fixation location and duration. Thus, the extra time spent fixating an informative object is the time needed to add the informative object to the schema for that scene. This research implies that information about the meaning of a scene is quickly acquired and guides further processing of the scene, including location of eye-movements, and fixation duration.

The results from Loftus and Mackworth's (1978) study support a schema hypothesis of object processing. Like a frame as outlined by Friedman (1979), a schema is a representation of the semantic category of a scene and contains information about what types of objects and their relations should be present in the scene. But, whether or not schematic information is guiding perceptual encoding is still not apparent. One concern centers around the dissimilarity of some of the informative objects from the noninformative objects in the scenes. Specifically, was there any difference in the physical characteristics between the informative objects and the noninformative objects and could this physical difference influence ordinal fixation number? This possibility was of some concern for the authors, resulting in removal of some scenes from the analyses. However, it is still unclear if this problem was present in the remaining scenes. For example, in the farm scene they describe, the octopus' physical characteristics are "squigglely" lines, arguably different from the linear lines of the tractor, and the barn, house and fence in the background. The octopus is obviously different from the other objects in the scene, and this attribute could be influencing where participants are directing their eye-movements.

Research by Metzger and Antes (1983) has addressed the availability of context and object information early in picture viewing and questions the findings of Loftus and Mackworth (1978). In this experiment, they examined the recognition accuracy for portions of a scene that contained object information (what they called high informative areas) with areas that contained context information (or medium and low informative areas) after presentation of a scene for either 10, 30, 50, 75, 100, 150, 300, or 1,000 ms. They reasoned that if object recognition mediated the development of context, high informative areas should have a greater recognition accuracy at the earlier presentation times than medium and low informative areas. This pattern would result because object information would have to be extracted very quickly (at the earlier presentation times) in order to be used to guide context development. If, on the other hand, context mediated order to be used to guide context development. If, on the other hand, context mediated object recognition, medium informative areas should have greater recognition accuracy at the earlier presentation times than high informative areas. In their study, participants were presented with pictures of scenes for one of the presentation times. These scenes had been divided into eight sections which had been rated by judges on the amount of information each section contained (high, medium, or low). After the scene had been presented, a visual mask was shown for 100 ms, followed by a target probe, which was one of the eight sections of the preceding scene. Participants then determined if the probe was from the stimulus picture.

Metzger and Antes (1983) found that for the 10-300 ms exposure durations, medium informativeness areas were recognized better than high or low informative areas, a result that contradicts the findings of Loftus and Mackworth (1978), and that all three area types were recognized equally well at the 1,000 ms exposure duration. They also found that the relative performance on high and medium informative areas was influenced by location of the target probe. When the probe occurred in peripheral locations, medium informative areas were better recognized than low or high informative areas. When the probe occurred centrally, high informative areas were recognized most accurately. Metzger and Antes concluded that areas in a scene that contribute to contextual information are recognized earlier than areas that rely to a greater degree on object recognition. Moreover, context information is available at exposure durations too short for an eye-movement to take place. However, they suggested that with the gradual improvement in recognition for all types of information over time, it is most likely that

While the operational definitions of informativeness may be different between the two studies, the findings of Metzger and Antes (1983) question those of the Loftus and Mackworth (1978) study. In the Loftus and Mackworth study, it was found that high informative *objects* were fixated earlier than noninformative *objects*. In the Metzger and Antes study, at exposure durations too short to make an eye-movement, medium informative areas (areas that do not convey any information about *objects* in the scene) were recognized more accurately than high informative areas (areas that *do* convey information about *objects* in the scene). Thus, in one study, object information appears to be extracted quickly and in the other, context information is extracted quickly. It is interesting to note that Metzger and Antes found better recognition of high informative areas in the central location of the scene; and while the location of the target object in all of the scenes used in the Loftus and Mackworth study is not known, in the example discussed, the target object does occupy the central location. Thus, it is possible that the high informativeness effect found by Loftus and Mackworth is confounded with location.

De Graef and colleagues posit a slightly different view of the locus of context effects on object processing (De Graef, Christiaens, & d'Ydewalle, 1990; De Graef, De Troy, & d'Ydewalle, 1992; De Graef, 1992). Using a paradigm where participants scanned a picture looking for non-objects while their eye movements were monitored, De Graef et al. (1990) found that first fixation durations (defined as the initial fixation on an object before any subsequent fixations, either on the same object or on another object) on objects in the scene did not differ for objects undergoing violations compared to "normal" objects in the scene. First fixation is posited to be a conservative measure of object encoding (Henderson, Pollatsek, & Rayner, 1989, cited in De Graef, et. al. 1990), meaning that it is less likely to reflect post-perceptual encoding processes. In this study, two objects were chosen to be target objects and were subjected to relational violations. These relational violations in the study were: size, position, support, and probability. Interposition was not used because of the inability to violate this relation without disturbing featural structure. De Graef et al. measured the first fixation durations on these target objects in the scene relative to other objects and found that context influenced object recognition, but that this influence was not always apparent on the first fixation on the scene. The lack of a violation effect on first fixation duration, then, suggested that context effects were not immediately present. However, fixations that occurred later during scene viewing (approximately 10 fixations later) appeared to be influenced by object violations, indicating that schematic influences occurred during later processing. This interaction of object violations and early versus late fixations suggests that initially scene context does not affect object perception, but later (after approximately 10 fixations) it does (Rayner and Pollatsek, 1992).

In summary, various studies using eye-movement monitoring tasks have shown that scene context information is apprehended very quickly and can influence object processing. Moreover, in some cases, scene context information appears to influence object encoding. Scene context information has been shown to influence mean durations of eye fixations on objects within a scene, memory for figurative detail of objects in scenes, and the general pattern of eye-movements while viewing a scene, i.e., how quickly a particular object will be fixated upon presentation of a scene. These results have led researchers to believe that scene context guides perceptual encoding. However, at least some studies have shown that when different measurements of eye-fixations are used, the pattern of results suggests that scene context influences post-identification processing.

The Object Detection Task:

The object detection task (Biederman, 1972) has been used to examine the effects of scene context on perceptual encoding. In the object detection task, the participant is given the name of an object that may or may not be present in a following scene. Next, a scene is briefly presented to the participant (for about 150 ms, or within the duration of a single eye fixation) followed by a mask. In the mask, there is a location marker, and the participant's task is to determine whether the target object appeared at the marked location. The dependent measure is the probability of correctly detecting the cued object as a function of whether it appeared in an organized or unorganized scene or in an appropriate or inappropriate scene context. Biederman and his colleagues posited that specific object-context relations (size, position, support, interposition and probability) influence perception of objects. This influence is believed to be the result of top-down processes guiding object encoding. According to Biederman, some of the influence is "semantic," having to do with the scene's meaning or referential content, while some is "syntactic," having to do with the scene's structure or organization. Biederman and his colleagues argued that syntactic and semantic information are extracted during the earliest eye fixations and not just from the foveal region of the visual field (Biederman, 1981; Biederman, Mezzanotte, & Rabinowitz, 1982). In their experiment, syntactic violations involved the relations of support and interposition. These relations refer to the idea that objects are supported by other objects and are occluded by objects that lie in front of them, respectively. Semantic violations involved the relations of probability, the likelihood that an object will occur in a scene; position, where that object is likely to occur in a scene; and size, the familiar size of objects.

In an experiment using the object detection task (Biederman et al., 1982), these relations were manipulated to test what type of information is extracted during the earliest fixations on a scene. Biederman and his colleagues were interested in the accessibility of these different types of relations, and specifically, if different types of relations are accessed before others. To examine this question, scenes were constructed so that objects in the scene underwent various relational violations and presented to participants using the object detection task. If different relations are accessed at different rates, as a bottom-up model of scene perception would predict, semantic violations would not influence object detection while syntactic violations would. This pattern is predicted because the bottom-up model of scene perception predicts that physical (syntactic) information is accessed faster than semantic information. However, Biederman and his colleagues point out that there is no guarantee that syntactic relations are accessed before semantic. To address the question of differential access of various relations, Biederman et al. violated multiple relations to examine the effects of multiple violations compared to single violations.

Biederman and his colleagues (1982) found that participants most accurately detected the presence of the target object when it occurred in its appropriate scene context and had not undergone any violations. Furthermore, the greater the number of violations a target object underwent, the less likely it was detected in a scene. Because scenes were presented for only 150 ms, it was posited that from a single fixation, schematic information about the scene is readily extracted. Biederman et al. concluded that during a single fixation, scenes can be identified, and various amounts of relational information about an object and the rest of the scene are obtained both at the fovea and in the periphery (Biederman et al. 1982). This information can then influence perceptual encoding of that object.

Other evidence that scene context influences perceptual encoding comes from Boyce, Pollatsek, and Rayner (1989). In a set of experiments, Boyce et al. were interested in the role scene backgrounds play in scene context effects and whether or not these effects were the result of global scene information or local-object information. According to the local-object priming hypothesis (Henderson et al., 1987), scene effects are the result of object to object priming. Thus, a target object that is related to the other objects in the scene will require less visual processing because it will be primed by the other objects in the scene. In Experiment 1, participants performed the object detection task with a scene presentation time of 150 ms. The scene contained either an episodically consistent background, an inconsistent background, or no background at all. Episodically consistent was defined as a scene containing objects that regularly co-occur in the real-world settings or environments. They found a significant interaction between background presence and consistency. Object detection was better in consistent backgrounds compared to their no background controls than in inconsistent backgrounds compared to their no background controls. Moreover, these effects were apparent in a 150 ms presentation of a scene, indicating that the meaning of the scenes was apprehended very rapidly.

In Experiment 2, Boyce et al. (1989) examined the effect the non-cued objects in the scene (the cohort set) had on detection of the target object, as a direct test of the local-object priming hypothesis. In this experiment, they manipulated the degree of relatedness to the target object of the non-cued objects in a scene, and whether or not the objects appeared in a consistent or inconsistent background as related to the target object or no background at all. Scenes were presented that had either a consistent background, an inconsistent background, or no background at all, and the target object either appeared with four related non-cued objects or four unrelated non-cued objects. They found that relatedness of non-cued objects to the target object did not influence detection, but whether or not the target object appeared in a consistent or inconsistent background did. This finding, they concluded, suggested that the effect of scene context arose from the level of the general background of the scene rather than from object to object priming as the local-object priming hypothesis predicted.

Concerned that a no background control may not be the appropriate control for the consistent and inconsistent background conditions, in Experiment 3, Boyce et al. (1989) compared object detection in consistent and inconsistent background scenes with a nonsense background scene condition that preserved background complexity but provided no real meaning. Again, they found that object detection was greater in consistent background scenes than in nonsense background controls. They concluded that episodically consistent backgrounds facilitated object encoding.

From these three experiments, Boyce et al. (1989) concluded that the facilitation of context comes from the global information conveyed in the background, and not local object information, and also that this global information is acquired early. Additionally, it would appear that object encoding and scene comprehension occur simultaneously and during the very initial fixations on a scene.

Boyce and Pollatsek (1992) also examined the effects of scene context on the identification of objects. In this study, they used a "wiggle" paradigm where an object in the scene was moved a short distance and then back to its original position. This paradigm, they argued, was an unobtrusive method of drawing the participant's attention to the target object and was somewhat natural in that the movement of objects in the environment often attracts visual attention to them. In their experiment, participants fixated a cross in the center of the screen. Next, a scene appeared on the screen and after 75 msec one of the objects (the target object) moved a small distance and returned to its original position. Participants fixated the target object and tried to name it as quickly as possible, with naming latency the dependent variable. Boyce and Pollatsek varied scene background to examine differences in identification of the "wiggled" object. Three types of scenes were created for each target object: consistent-background scenes, inconsistent-background scenes, and nonsense background scenes. The inconsistent-background scenes were created by switching the non-cued objects in one scene with the non-cued objects from a paired scene. Nonsense background scenes similar to those used in Boyce et al. (1989) were created for both consistent and inconsistent-background scenes to control for the effect of different object locations between the two. Boyce and Pollatsek found that the “wiggled” object was named faster in the consistent-background condition than the inconsistent-background condition and concluded that scene context information was acquired early and did affect perceptual encoding. Furthermore, consistent context facilitated perceptual encoding and inconsistent context inhibited perceptual encoding.

Criticisms of the prior experimental paradigms.

While previous research seems to indicate that scene context influences and possibly even guides perceptual encoding, there are some concerns with the paradigms used to explore these questions and thus used to draw these conclusions. A review of some of the criticisms leveled against prior experimental paradigms follows.

In the object detection task, one concern is the potential for participants to generate a guessing strategy to perform the task. As stated before, in this paradigm, participants are given the name of the target object before they are shown a scene, and herein lies the problem. Giving the name of the object before the scene is presented potentially allows the participant time to generate certain expectations about where the named object may or may not occur. For example, given the name "couch" as the target object, a participant not only knows what to look for in a scene, but to some extent, where to look for it within a scene. If the participant knows where to look (e.g., couches usually are on the floor and thus should appear low in the scene), then when the couch is floating somewhere inappropriately high, participants are likely to miss it because they will be more likely be attending to the lower part of the scene. So, this circumstance predicts poorer performance when position is violated.

Another question concerning the object detection task is whether or not object detection is the same as object identification. By giving participants the name of the target object, they can look for particular features of that object that may aid in detecting the presence or absence of the target object in the scene. The question then, is whether this type of processing is different from situations where the features of an unknown object are matched against the features of a stored memory trace, the process that more likely occurs during normal scene viewing and object processing.

Undoubtedly, eye-movement monitoring provides a fairly natural and unobtrusive method for examining scene processing. However, while eye-movement experiments have demonstrated that participants fixate objects that "belong" in a scene for less time than objects that do not, it is unclear what cognitive processes different measures of fixation time reflect. The particular question of interest is whether or not fixation duration measures used in previous experiments are reflecting perceptual encoding of objects or post-identification processes. One problem, Henderson (1992) points out, concerns the lack of a demarcation between identification and further post-identification processes in the literature. Gaze duration (defined as the time of all initial fixations on an object prior to leaving that object for the first time, including other intra-object fixations) has been used by a number of studies demonstrating context effects on eye-movement patterns (Friedman, 1979; Loftus and Mackworth, 1978). Because this type of measure reflects a fairly general amount of processing time, it is likely that this measure reflects processes occurring after the object has been identified. Consequently, it is unclear whether or not the context effects obtained in these experiments reflect object encoding or some other post-perceptual processing, for example, memory encoding. Additionally, Henderson points out that eye movement patterns may change based on the viewing task and thus these findings may reflect strategies used by participants to perform the tasks of the experiment. These concerns are readily apparent given the inconsistency of the results from different studies using eye-movement paradigms (e.g., Loftus & Mackworth, 1978; De Graef et al., 1990).

Finally, several experiments showing scene context effects have used object naming as the dependent measure (Boyce et al., 1991), and it is possible that naming times reflect processing that occurs after the object has been identified.

In all, these criticisms illuminate the fact that previous research paradigms leave open the possibility that it is not perceptual encoding processes that are affected by scene context but post-perceptual encoding processes. As a result, the effects of scene context on object processing are not fully understood. The perceptual schema hypothesis posits that context effects result from a top-down influence on perceptual encoding, but it is now apparent that experiments used to formulate this hypothesis, and to draw this conclusion, may have been flawed in some manner. What appears to be needed, then, is a task that mimics normal scene viewing without the concern that the dependent measure reflects post-identification processes.

Evidence for different types of information in complex scenes.

Mandler and Johnson (1976) describe four types of information in complex visual scenes. This taxonomy of information includes: inventory information, which specifies the objects in the picture; spatial location information, which specifies the location of the objects in the picture, including relative location to other objects; descriptive information, which specifies the figurative detail of the objects; and spatial composition information, which specifies areas of filled or empty spaces and the density of filled spaces. Concerned with the encoding and storage of information and the effects that schemata have on this process, Mandler and Johnson used recognition memory tests to assess long-term retention of these four types of visual information from a scene. By manipulating the types of information in the scenes along with their organization, they examined participants' memory for visual information. Participants were shown a sequence of 10 pictures for five, 20, or 60 sec presentation times followed by a same/different recognition test. During the recognition test, participants were shown a sequence of 100 pictures and were instructed to determine which of the 100 pictures were ones that they had seen earlier. They found that spatial location information was better recognized in organized scenes, while spatial composition was better recognized in unorganized scenes, representing schemata-driven and non-schemata-driven processing, respectively. Descriptive and inventory information proved to be independent of picture organization. Mandler and Johnson concluded that there is differential memory for different kinds of visual information. Moreover, schemata have varying effects on the influence encoding and storage of the different types of visual information, aiding memory of spatial location information, while inhibiting memory of spatial composition information.

In a follow-up study, Mandler and Ritchey (1977), using the same taxonomy of visual information as Mandler and Johnson (1976), examined memory for visual information over extended periods of time. Participants were shown a sequence of eight pictures and tested for recognition memory either "immediately" after study, or after retention intervals of one day, one week, or four months. Like Mandler and Johnson, Mandler and Ritchey also found marked differences in memory for different types of visual information. Spatial relation information was better retained in organized pictures and this information persisted over the 4 month period. Descriptive information, on the other hand, while independent of organization, was not well retained over extended periods of time. Inventory information was found to be influenced by picture organization, but lasted over the 4 month period. However, unlike Mandler and Johnson, Mandler and Ritchey found spatial composition information to be poorly retained, even at the immediate testing interval. They concluded that the memory representation of scenes contains information regarding inventory of objects, and their relative locations, but not descriptive or spatial composition information. Additionally, as time progresses, recognition memory is more schemata-driven than immediately after encoding.

These studies suggest that Mandler and Johnson's (1976) taxonomy of visual information is a viable description of at least some of the types of information encoded from scenes. However, while this taxonomy has gained support from studies examining recognition memory following relatively long scene presentations, it remains to be seen whether or not it will explain what type of information is encoded from initial eye fixations.

Evidence for rapid scene comprehension.

Some research has addressed scene recognition performance following relatively brief scene presentations. Potter and Levy (1969) examined memory for visual information following presentation at or around the time of a single eye fixation. Using presentation rates between 125 and 333 msec and a yes/no recognition test, Potter and Levy found memory for pictures to be greater at the longer presentation times. From these findings, they concluded that the recognition memory for pictures presented in sequence is dependent on the amount of time the individual picture is in view and that processing occurs until there is a change in visual stimuli.

In later work, Potter (1976) again examined recognition memory for rapidly presented pictures. Using a target search paradigm, Potter tried to determine whether rapidly presented pictures (presentation rates of 113, 167, 250, and 333 msec) are identified and then forgotten, or not identified at all. Some participants searched for a target picture in a series of pictures, having been given a brief title for the picture, while others were first shown the target picture itself. A third group of participants performed a yes/no recognition test following each of eight sequences of pictures. At presentation times of 113 msec and above, detection of the target given either picture or title preview was above chance. Recognition memory performance, however, was above chance only at the longer presentation rate. Thus, while pictures may be identified at very short presentation rates, recognition memory for these pictures at these rates was very poor. Potter theorized that the time between identification and recognition memory must be a time of consolidation, and posited that this consolidation occurs in a short-term conceptual memory. Additional research found that a mask presented between the first scene and the test scene that allowed at least 300 ms of processing time after a short first scene presentation improved recognition memory performance. Potter posited that the mask stopped visual processing of the first scene but allowed conceptual processing of the scene to occur, thus improving recognition memory performance. Most importantly for the present experiments, it appears that participants can perform recognition tasks following relatively brief scene presentations.

Work by Intraub (1980, 1984) supports the idea that following brief presentations, additional processing time is needed to perform above chance on a recognition task. Intraub examined the possibility that pictorial encoding involves a process whereby memory representation increases over time, and that this process extends beyond the duration of the stimulus. According to Intraub, one possible hypothesis is that picture processing is an all-or-nothing process that requires a fixed amount of time. Under this theory, participants can attend to a picture until it has been encoded, and as the I.S.I. is decreased, they will miss pictures that are presented while they are encoding a previous picture. On the other hand, if picture processing is a continuous process, then as I.S.I. is decreased, fewer details of the picture will be encoded. Recognizing that a normal recognition test would not address this question (a minimal amount of semantic or visual information may be sufficient for a recognition response), Intraub introduced a task whereby participants had to choose mirror reversals of the target on half of the trials in the recognition task. In this experiment, participants were shown a picture either for 5 sec with no I.S.I. or for 110 msec with an I.S.I. of 4890; 1390; 620; 385; or 0 msec. She found that reversing the picture did not affect participants' ability to tell that the picture had been seen before, but that the ability to tell that a correctly identified picture was mirror reversed decreased not only with stimulus duration, but also when stimulus duration was held constant and the I.S.I. was reduced. Intraub posited that the I.S.I. used in the experiments allowed time for memory representation encoding following brief presentations. Consequently, when there is no I.S.I., the following picture interferes with the encoding process of the previous picture. Intraub concluded that visual information encoding extends beyond the duration of the stimulus and is somewhat independent of the number of eye fixations. These results have been taken to support Potter's (1976) hypothesis that visual information encoding extends beyond the physical duration of the stimulus.

The present experiments.

While previous research has examined the taxonomy of visual information in memory, as well as recognition memory for rapidly presented visual stimuli, none has addressed the combination of the two areas. From previous research, it is evident that scene context information is extracted very quickly (Intraub, 1980, 1984; Potter, 1976; Boyce et al., 1989, 1992; Biederman, 1982). Also, there appear to be different categories of visual information in long-term memory (Mandler and Johnson, 1976; Mandler and Ritchey, 1977). However, there has been no empirical research that has examined how these different categories of visual information might be extracted initially from a scene and in particular, what effects scene context has on these processes. In other words, will scene context influence the detection of different types of changes to objects in briefly presented scenes, indicating that it influences perceptual encoding of that object?

In the reported set of experiments, context effects on perceptual encoding were examined using a new paradigm. In this paradigm, participants were presented with a study scene, followed by a mask, and then a test scene which was the same scene with one object changed or not. Their task was to determine whether the study and the test scenes were identical. On trials in which the two scenes differed, a target object underwent a particular type of transformation between the presentations of the two scenes. Scene context was manipulated by switching target objects across paired scenes, creating appropriate and inappropriate scene context conditions. The duration of the initial study scene was varied across experiments. By using brief presentation times for the study scene (250 ms, or approximately the time of a single eye fixation) in an initial experiment, the effects of scene context on perceptual encoding was examined. In subsequent experiments longer study scene durations were used to examine the influence of scene context on potential post-perceptual encoding processes. Participants were presented with the study scene, followed by a noise-filled mask for 400 msec. Because it did not require conceptual processing itself, this type of mask stopped visual processing of the scene after the initial presentation, but allowed conceptual processes to continue. As a result, with a 400 msec visual mask, participants should be able to process the first scene sufficiently to perform the same-different task. Following the mask, participants were presented with either the picture of the same scene or one in which the target object had undergone a particular transformation in the test scene, at which time they performed a same/different recognition test. The major advantage of this same/different task over previous experimental paradigms is that there is no decision component in terms of identifying a target object, its location, or naming the target object. In this task, participants are simply presented with one fixation of a scene and then determine whether any visual information has changed between this study scene and a subsequent test scene. If participants notice the change, they must have encoded the information. Accuracy and sensitivity (d') in detecting the manipulations of the target object as a function of scene context were examined.

In the experiments presented here, the question of interest was whether there would be differential effects on the type of visual information encoded about a particular object based on scene context. To address this question, scenes were paired together and scene context was manipulated by placing a target object from one scene into its paired scene, producing the inappropriate scene context condition. This manipulation should determine whether or not there is differential encoding of the types of information the manipulations address about the target object as a function of scene context.

To examine the different types of information claimed to be extracted from a scene, two target object manipulations were used: deletion and orientation change. The deletion manipulation is believed to address the degree to which the object is encoded at all. The orientation change manipulation is posited to address the degree to which specific visual characteristics have been encoded.

While any given experiment will only use one presentation time for the study scene, together the experiments will address the time course of visual information encoding by varying the amount of time the study scene is presented across different experiments. Interactions between presentation time, scene context, and target object manipulation may lend support to the hypothesis that schemata do not influence perceptual encoding but post-perceptual encoding.

Experiment 1.

The purpose of Experiment 1 was to validate the scene context manipulation to be used in the subsequent experiments. In this experiment, participants were shown the appropriate and the inappropriate scene conditions for the 24 scenes used in the following experiments and asked to determine if all of the objects fit in the scene.

Method

Participants: Sixteen introductory psychology students at Michigan State University participated in this experiment. Participants received partial credit for their introductory psychology courses for participating in this experiment. They received 1 credit for every half hour of experiment participation. All participants had normal or corrected to normal vision.

Apparatus. Responses from participants determining whether or not all of the objects fit in the scene were recorded by a 486-66 PC microcomputer. Participants used a button-box interfaced with the computer to start each trial and to report whether all of the objects fit in the scene.

Materials. Twenty-four scenes constructed by De Graef (1990) were used. De Graef created the scenes by taking photographs of real scenes in the environment. He then created slides of each of the 24 natural photographs, and projected them onto a screen and drew line drawings of the scenes which were used in his studies. In the present study, target objects were selected from each of the scenes, and then scenes were paired together so that the target object, when placed in its paired scene, would create a scene in which the target objects did not fit the scene (the inappropriate scene context condition). Target objects for paired scenes were matched for general size and shape and occupied the same location in the scene when placed in the paired scene (see Appendix A).

Participants only saw the 24 scene examples in which the target object was appropriate (the appropriate scene context condition) and the 24 scene examples in which the target object was inappropriate (the inappropriate scene context condition) and none of the scenes in which the target object had been manipulated.

Procedure. Participants were tested individually. Upon arrival at the experimental session, participants were seated in front of the computer and button box. The experimenter then explained to them that in this experiment, they would be presented with a scene and their task would be simply to determine if all of the objects in the scene fit in the scene or not. If they believed that all of the objects fit the scene, they were to press the "yes" button, and if they believed that one or more of the objects did not fit the scene for any reason to press the "no" button. The experimenter then answered any questions participants had about the task before they began.

During each trial of the experiment, the participant was presented with a fixation cross at which they were to direct their gaze. They pressed a button on the button box and the scene was presented on the computer monitor. Participants were free to examine the scene at their own pace and when they were ready to make a decision, they pressed either the "yes" or "no" button, at which time the scene disappeared from the computer monitor. Participants then pressed a button again to start the next trial. Each participant saw all 48 scenes in a completely randomized order. After participants completed the 48 trials, they were debriefed by the experimenter and thanked for their participation. The entire session lasted approximately 15 minutes.

Results and Discussion. An ANOVA was conducted on the percentage "yes" responses for the appropriate and inappropriate scene context condition scenes. There was a main effect of scene context, $F(1,15)=779.18$, $MSe=.01$, $p<.005$. Participants responded "yes" 89.3 % of the time for the appropriate scene context condition, and "no" 89.5% of the time in the inappropriate scene context condition.

This task was an attempt to test the context manipulation used in the following experiments, and in particular, if the target object fit in the appropriate scene context condition and did not fit in the inappropriate scene context condition. In this experiment, participants were asked to determine if all of the objects in the scene "fit" in the scene. From the results in the experiment, it appears that participants believed that all the objects fit in the appropriate scene condition, and that at least one object did not fit in the inappropriate scene context condition. Because the only object that changed in the inappropriate scene context condition was the target object, it can be concluded that in the inappropriate context scenes, participants were basing their decision on the inappropriate target object. As a result, it would appear that scene context was manipulated adequately in the construction of the scenes used in the following experiments.

Experiment 2

The purpose of Experiment 2 was to examine the effects of scene context on the perceptual encoding of objects within that scene. Participants viewed a study scene for 250 ms, followed by a pattern mask for 400 ms, followed by the test scene with the target object changed or not. The participant's task was to press a button to indicate whether the two scenes were the same. Two types of target object manipulations were examined: object deletion and object orientation reversal. If scene context facilitates the encoding of the presence of a consistent object, then deletion detection should be better when the object fits than when it does not fit in the scene. If context facilitates the encoding of spatial information, then orientation change detection should be better when the object fits than when it does not.

Concerning scene organization, Biederman and his colleagues (1981) found relational violation effects on object identification; and while the present experiment will not manipulate context in this manner, it does seem reasonable to assume that there may be differential encoding of visual information as a function of scene context. Thus, if the schema hypothesis is correct, participants should notice target object manipulations more in the appropriate context condition in this experiment. No differences in information encoding between the two types of scenes - appropriate and inappropriate - could mean that scene context effects do not occur during perceptual encoding but are post-perceptual in nature.

Method

Participants. Twenty-four introductory psychology students at Michigan State University participated as participants in this experiment. Participants received partial credit for their introductory psychology courses for participating in this experiment. All participants had normal or corrected to normal vision and were naive with respect to the purpose of the experiment. None of the participants had participated in Experiment 1.

Apparatus. Response times and accuracy to determine if the two scenes were the same were recorded by a 486-66 PC microcomputer with a NEC XE15 (Multisync) VGA monitor. Participants used a button-box interfaced with the computer to start each trial and to make their yes/no decisions.

Materials. The 24 appropriate context and inappropriate context scenes used in Experiment 1 were used in this experiment. For each scene in both the appropriate and inappropriate context conditions, target objects were manipulated in one of two ways to create the different scene conditions. For the deletion condition, the target object was removed from the scene. For the orientation change condition, the target object was rotated about its vertical axis. In the same condition, the target object underwent no type of manipulation. Paper copies of the three types of each of the 24 scenes were scanned into the computer for later presentation to participants on a computer monitor. On the monitor, the scenes subtended a visual angle of 23.8 degrees (width) by 17.7 degrees (height).

The mask used between presentations of the first scene and the test scene was constructed by superimposing the computer scanned images of the scenes on top of each other. Once all 24 scenes were superimposed, the resulting image was then flipped on its horizontal axis, and this image was superimposed on top of the original superimposed image. This created a mask where no given object from any of the scenes was discernable. The size of the mask was the same as the size of the scenes.

Procedure. Participants were tested individually. Upon arrival at the experimental session, participants were seated in front of computer monitor, button box, and head and chin rest. The experimenter then explained to them that in this experiment, their task would be to determine if two scenes presented to them on the computer monitor were the same or different. Because it was important that every participant remain the same distance from the computer monitor, they were also told that the head and chin rest would be used in the study and the height of the chair was adjusted to insure comfort in the head and chin rest.

During each trial of the experiment, participants saw a prompt instructing them to press a pacing button on the button box to begin the trial. When the participant pressed the button, a fixation cross remained at the center of the screen for 500 ms followed by the presentation of the first scene for 250 ms. After the first scene was presented, the mask was presented to the participant for 400 ms, followed by the test scene which was either the same scene as the first one during the trial, or the same scene with a manipulation of the target object (see Figure 1). Participants either pressed the left button for "same" or the right button for "different." After they made a decision, they were again to press the pacing button on the button box to begin the next trial.

Insert Figure 1 here

Before participants started the experimental session, the experimenter showed them an example of each of the types of manipulations that could occur for a given scene. This demonstration used a scene that was not part of the experiment. The experimenter explained to them that the manipulation could occur for any object in the scene and that the manipulated object could occur at any location in the scene. Next, participants were run in a practice block of 16 trials (2 scenes X 2 scene context conditions X 4 target object manipulations). 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 participants had about the procedure and then the participant proceeded to the 192 experimental trials. The participants completed the experimental trials without the experimenter being present in the participant running room. After the experiment, participants were debriefed by the experimenter and thanked for their participation in the experiment. The entire session lasted about 35 minutes.

Design and Analysis. The design of the experiment was a 2 (scene context: appropriate, inappropriate) X 4 (target object manipulation: same, same, deletion, orientation change) factorial. There were two levels of the same condition to equate the number of "same" trials with the number of "different" trials. Scene context and target object manipulation were both within participant variables. All participants saw all 192 trials which were completely randomized. An Omnibus ANOVA that included all factors was conducted on mean percentage correct and d' response data.

Results. Mean percentages of correct responses are shown in Table 1. There was no main effect of scene context, $F(1,23)=.02$, $MSe=.01$, $p>.05$. Participants responded correctly 60.0% of the time in the appropriate scene context condition and 57.4% of the time in the inappropriate scene context condition. However, there was a significant main effect of target object manipulation, $F(3,23)=21.03$, $MSe=.04$, $p<.005$. When the study and the test scene were the same, participants responded correctly 75.9% of the time. In the deletion condition, participants responded correctly 31.2% of the time. In the orientation change condition, participants responded correctly 51.7% of the time. There was no significant interaction between scene context and target object manipulation,

$F(23,69)=.58$, $MSe=.01$, $p>.05$.

Insert Table 1 here

An additional ANOVA was conducted on participants' mean d' data to determine if participants were detecting the target object manipulations differentially based on whether the target object appeared in the appropriate or inappropriate scene context (see Table 1). Mean d's for the deletion and orientation change condition were .30 and .92 in the appropriate scene context, respectively and .19 and .73 in the inappropriate scene context, respectively. There was a marginal main effect of context, $F(1,23)=3.07$, $MSe=.16$, $p=.07$. As with the percentage correct data, there was a significant main effect of target object manipulation, $F(1,23)=61.58$, $MSe=.13$, $p<.05$. Participants noticed the orientation change manipulation significantly more than the deletion manipulation in both the appropriate and inappropriate scene contexts. Finally, there was no interaction between scene context and target object manipulation, $F(1,23)=.75$, $MSe=.07$, $p>.05$. Participants did not differentially detect the target object manipulation as a function of whether the object appeared in the appropriate or inappropriate scene context.

Because of the concern about participants becoming aware that the scenes were being repeated several times during the experimental session, a quartile analysis was conducted to examine participants' performance at the beginning and the end of the experimental session. Mean percentage correct responses for the first and fourth quarter of the trials are shown in Table 2. There were no significant differences in correct responses or interactions between the first and fourth quarter of experimental trials (all F's < 1).

Insert Table 2 here

Discussion The data from Experiment 2 indicate that participants may have differentially encoded information about a target object as a function of scene context during initial fixation on a scene. In this experiment, accuracy to detect differences between the first and second scenes was slightly better in the appropriate context condition than in the inappropriate condition, as indicated by the marginal main effect of scene context in the d' data. However, participants did not encode differently the target object's orientation or presence relative to whether or not the target object appeared in the appropriate condition or the inappropriate condition. When the target object was deleted, participants were just as accurate in noticing this change when the object occurred in the appropriate scene context as when it occurred in the inappropriate scene context. Interestingly, participants were more accurate at noticing orientation changes than deletions, but nevertheless they encoded these changes equally when they occurred in the appropriate or inappropriate scene context. The marginal effect of context in the d' analysis does indicate that participants were slightly more sensitive to the target object manipulations in the appropriate scene context condition than in the inappropriate scene context condition. Taken as such, the results offer, at best, weak support for a schemata hypothesis as outlined by Boyce et al. (1989, 1992) and Biederman (1981).

The lack of a significant scene context effect on perceptual encoding in Experiment 1 may be the result of scene context not being able to influence initial perceptual encoding, or because scene interpretation requires more than 250 msec. Again, the d' data suggest a slight context effect. To examine this possibility, study scene presentation time was increased in Experiment 3 from 250 ms to 500 ms. Because Biederman et al. (1982) found context effects at 150 ms, a presentation duration of 500 ms in this experiment seems to be sufficient to uncover scene context effects in this paradigm if they exist.

Experiment 3.

In Experiment 2, there was a marginal effect of scene context on perceptual encoding as predicted by the perceptual schema hypothesis. Experiment 3 attempted to find a reliable context effect by giving the participant a longer period of time to examine the study scene in the trial. By increasing this presentation time, participants could make more than one eye fixation, though eye-movements were not recorded. Also, if the marginal effect of scene context found in the d' analysis in Experiment 2 reflects a true effect of context on encoding, it should replicate in Experiment 3 where more time is available for deriving the meaning of the scene.

Method

Twenty-four introductory psychology students at Michigan State University participated in this experiment. Participants received partial credit in their introductory psychology courses for participating in this experiment. All participants had normal or corrected to normal vision and were naive with respect to the purpose of the experiment. The apparatus, materials, procedure, design and analyses were the same as in Experiment

2 with the exception that the first scene was presented for 500 ms instead of 250 ms.

Results. Mean correct response percentages and d's are shown in Table 3. There was a significant main effect of context, $F(1,23)=5.36$, $MSe=.01$, $p<.05$. Participants responded correctly 59.7% of the time in the appropriate context condition and 57.1% of the time in the inappropriate context condition. Like Experiment 1, there was a significant main effect of target object manipulation, $F(3,23)=256.55$, $MSe=.01$, $p<.005$. When the study and the test scene presentations were the same, participants responded correctly 80.0% of the time. When the test scene contained a deletion, participants responded correctly 26.5% of the time. When the test scene contained an orientation change of the target object, participants responded correctly 47.2% of the time. Again, participants were more accurate at noticing an object when it had been mirror-reversed than when it had been deleted from the scene. Finally, the interaction between context and target object manipulation was again not significant, $F(3,69)=1.68$, $MSe=.01$, $p>.05$.

Insert Table 3 here

Again, an additional ANOVA was conducted on participants' mean d' data to determine if participants were detecting the target object manipulations differentially based on whether the target object appeared in the appropriate or inappropriate scene context (see Table 3). Mean d's for the deletion and orientation change condition were .32 and .96 in the appropriate scene context, respectively and .23 and .76 in the inappropriate scene context, respectively. There was no main effect of context, $F(1,23)=2.55$, $MSe=.20$, p>.05. As with the percentage correct data, there was a significant main effect of target object manipulation, $F(1,23)=67.61$, $MSe=.12$, $p<.05$. Participants noticed the orientation change manipulation significantly more than the deletion manipulation in both the appropriate and inappropriate scene contexts. Finally, there was no interaction between scene context and target object manipulation, $F(1,23)=1.03$, $MSe=.07$, $p>.05$. Again, participants did not differentially detect the target object manipulation as a function of whether the object appeared in the appropriate or inappropriate scene context.

To check for practice effects, a quartile analysis was again conducted on the first quarter and the last quarter of experimental trials of Experiment 3. Mean percentages of correct responses are shown in Table 4. Like Experiment 2, there were no significant differences between participants' percentage correct responses between the first and fourth quarter and no interactions, F's <1.

Insert Table 4 here

Discussion. In Experiment 3, finding a main effect of scene context in the percent correct data provides some evidence for there being an influence of scene context on object encoding. Participants were more accurate in responding in the appropriate scene context than they were in the inappropriate scene context condition. Participants did not, however, differentially encode target object information as a function of whether the target object appeared in the appropriate scene context or the inappropriate scene context. Finally, the same counterintuitive finding of participants noticing orientation changes more than deletions also occurred.

Once more, the data from the d' analyses failed to support a perceptual schema hypothesis of object processing. However, the main effect of scene context in the percent correct data suggests that scene context may have some type of effect on perceptual encoding. When participants were given more time and the opportunity to make two or maybe three eye fixations, they did not encode target object information any differently when the object occurred in its appropriate scene than when it occurred in an inappropriate scene. Thus, given more time to process the scenes, it appears that scene context does not influence initial perceptual encoding.

While contrary to the assumption that the gist of a scene is rapidly apprehended as the perceptual schema hypothesis claims, one impetus for doing the present experiment was concern that the scene interpretation requires slightly more than 250 msec. This concern seems highly unlikely given the reported results from previous research (i.e., Biederman et al., 1982; Boyce et al., 1989). Finding the main effect of scene context with a 500 msec first scene presentation time supports a view that scene context may influence more post-perceptual encoding. As a result, Experiment 4 was an attempt to examine scene context effects on post-perceptual encoding using an even longer first scene presentation time.

Experiment 4

De Graef et al (1990) used a set of stimuli very similar to those used here and found that scene context influenced post-perceptual encoding. In their study, participants viewed scenes while their eye movements were recorded. The participant's task was to count the number of non-objects in each scene. The relationship of the target objects to the scene context was manipulated. De Graef et al. found that scene context influenced first fixation duration on the target object when the target object was fixated around the tenth eye fixation on the scene and concluded that scene context does not influence initial perceptual encoding. The pattern of data reported by De Graef et al. seems to indicate that if scene context influences object encoding, it does not do so during the first one to two eye fixations on the scene but may do so at some time after about ten fixations on the scene. Experiment 4 was a replication of Experiments 2 and 3 reported here with the study scene's presentation duration increased from 500 ms to 2500 ms. This presentation duration approximated the amount of time De Graef et al. found was needed before scene context effects were uncovered in his experiment. With this presentation duration, participants were allowed to make as many as ten eye fixations, which may allow scene context information to influence object processing. Furthermore, with this presentation duration, the deletion manipulation may be detected more by participants. However, at this presentation time, it is more likely that post-perceptual encoding is what is being influenced by scene context.

Method

Twenty-four introductory psychology students at Michigan State University participated in this experiment. Participants received partial credit for their introductory psychology courses for participating. All participants had normal or corrected to normal vision and were naive with respect to the purpose of the experiment. The apparatus, materials, procedure, design and analyses were the same as in Experiments 2 and 3 with the exception that the first scene was presented for 2500 ms instead of 250 or 500 ms.

Results. Mean correct response percentages and d's are shown in Table 5. There was no main effect of context, $F(1,23)=1.47$, $MSe=.005$, $p>.05$. Participants responded correctly 60.1% of the time in the appropriate context condition and 58.8% of the time in the inappropriate context condition. Like Experiments 2 and 3, there was a significant main effect of target object manipulation, $F(3,23)=456.66$, $MSe=.01$, $p<.005$. When the study and the test scene presentations were the same, participants responded correctly 87.8% of the time. When the test scene contained a deletion, participants responded correctly 15.0% of the time. When the test scene contained an orientation change of the target object, participants responded correctly 47.1% of the time. Again, participants were more accurate at noticing an object when its orientation had been changed than when it had been deleted from the scene. Finally, the interaction between context and target object manipulation was again not significant, $F(3,69)=1.85$, $MSe=.004$, $p>.05$.

Insert Table 5 here

As with the previous experiments, an additional ANOVA was conducted on participants' mean d' data to determine if participants were detecting the target object manipulations differentially based on whether the target object appeared in the appropriate or inappropriate scene context (see Table 5). Mean d's for the deletion and orientation change conditions were .14 and 1.26 in the appropriate scene context, respectively and .18 and 1.12 in the inappropriate scene context, respectively. There was no main effect of context, $F(1,23)=.69$, $MSe=.16$, $p>.05$. As with the percentage correct data, there was a significant main effect of target object manipulation, $F(1,23)=250.40$, $MSe=.10$, $p<.05$. Participants noticed the orientation change manipulation significantly more than the deletion manipulation in both the appropriate and inappropriate scene contexts. Finally, there was an interaction between scene context and target object manipulation, $F(1,23)=4.47$, $MSe=.06$, $p<.05$. Participants differentially detected the target object manipulation as a function of whether the object appeared in the appropriate or inappropriate scene context. Specifically, detection of the orientation change manipulation was greater in the appropriate context than in the inappropriate context. Detection of the deletion manipulation was the same in both the appropriate and the inappropriate scene context.

To check for practice effects, a quartile analysis was again conducted on the first quarter and the last quarter of experimental trials of Experiment 4. Mean percentages of correct responses are shown in Table 6. Like Experiments 2 and 3, there were no significant differences between participants' percentage correct responses between the first and fourth quarter and no interactions, F's <1.

Insert Table 6 here

Discussion. Like Experiments 2 and 3, the percent accuracy data did not provide evidence for there being an influence of context on object encoding at 2500 msec. However, the d' data showed that at 2500 msec, there is an interaction of context and the target object manipulations used in these experiments. Participants detected the orientation change better when the target object appeared in the appropriate scene context than in the inappropriate scene context. However, like the previous experiments, detection of the deletion was the same in both scene contexts.

The fact that the scene context did not influence processing of the target object manipulations until the study scene was presented for 2500 msec provides some support for the hypothesis that scene context does not influence perceptual encoding. De Graef and his colleagues (1990) found that when a more conservative measure of encoding time (first fixation) was used, scene context did not influence eye-movement patterns until around the tenth eye fixation on the scene. From these findings, they concluded that scene context influences post-perceptual encoding. However the main effects of scene context in the earlier experiments reported here suggests that some effects on perceptual encoding may exist.

In the three experiments reported here, scene context influenced object processing at the earlier study scene presentation times, but did have an influence on the detection of the target object manipulations until the study scene was presented for a longer period of time. Taken together, these results can support the view that scene context guides perceptual encoding, as proposed by the perceptual schema hypothesis. However, it would also appear that scene context information affects post-perceptual encoding, such as memory or meaning encoding, for example.

Subsidiary Analyses, Experiments 1-4

Examination of the item data from Experiment 1 revealed that the scene context manipulation was stronger for some scenes than for others. As a result of this finding, additional ANOVAs on the mean percentage correct and d' data were conducted on the participant data from Experiments 2, 3, and 4 with "goodness of scene" as an additional factor to see if scene context would influence perceptual encoding for the better scenes. "Goodness of scene" was defined as a median split of the 12 best scenes and the 12 remaining scenes as indicated by participants responses in Experiment 1 (see Table 7). The 24 scenes were divided in the following way. Participants' mean percentage of "no" responses to the question "Do all the objects in the scene "fit" were tabulated for the 24 appropriate scenes and the 24 inappropriate scenes. Next, the proportion of times participants said "no" for the appropriate scenes was subtracted from the proportion of times participants said "no" for the inappropriate scenes. This difference for each scene was then rank ordered, with the 12 largest differences constituting the 12 "best" scenes.

Insert Table 7 here

Data from the median split analysis of Experiment 2 are shown in Table 8. In Experiment 2, in the percent correct data, there was no main effect of "goodness of scene," $F(1,23)=1.91$, $MSe=.0108$, $p>.05$. Participants' accuracy was 57.9% in the 12 "best" scenes and 59.4% in the 12 remaining scenes. There was an interaction between "goodness of scene" and scene context, $F(1,23)=4.10$ $MSe=.0170$, $p<.05$. Responses were 60.6% and 55.2% accurate for the appropriate context and the inappropriate context conditions respectively in the "best" scenes and 59.4% and 59.4% correct for the remaining scenes. There was an interaction between "goodness of scene" and target object manipulation, $F(3,69)=3.48$, $MSe=.0122$, $p<.05$. In both the 12 "best" scenes and the 12 remaining scenes, participants responded more accurately in the same condition than in either the deletion or the orientation change conditions, 75.5%, 32.4% and 48.3%, respectively for the 12 "best" scenes and 76.7%, 30.0% and 55.0%, respectively for the 12 remaining scenes. Finally, there was an 3-way interaction of goodness of scene, scene context, and target object manipulation, $F(3,69)=3.35$, $MSe=.0106$, $p<.05$. In the d' data, there was a main effect of scene context in the 12 best scenes, $F(1,23)=12.04$, $MSe=.302$, $p<.05$. Mean d' was .403 in the appropriate scene context and .313 in the inappropriate scene context condition. There was also a main effect of target object manipulation, $F(1,23)=30.92$, $MSe=.166$, $p<.05$. Mean d' was .277 and .739 for the deletion and orientation change condition, respectively. Finally, there was also a significant interaction of scene context and target object manipulation, $F(1,23)=5.026$, $MSe=.132$, $p<.05$. Mean d' was .388 and 1.017 for the deletion and the orientation change conditions, respectively in the appropriate scene context and .165 and .461 for the same conditions in the inappropriate scene context condition. In the 12 remaining scenes, there was no main effect of scene context, $F(1,23)=.264$, $MSe=.627$, $p>.05$. There was a main effect of target object manipulation, however, $F(1,23)=43.340$, $MSe=.309$, $p<.05$. There was, however, no interaction $F(1,23)=.185$, $MSe=.309$, $p>.05$. So it appears that scene context does influence object encoding when the scenes are separated by strength of the context manipulation of the scenes.

Insert Table 8 here

In Experiment 3, the additional mean percentage correct and d' ANOVAs done with goodness of scene included in the analysis revealed the following (see Table 9). In the percent correct data, there was a marginal main effect of goodness of scene, $F(1,23)=3.04$, $MSe=.0113$, $p=.09$. Participants' accuracy was 57.6% in the 12 "best" scenes and 59.5% in the remaining scenes. There was a main effect of scene context, $F(1,23)=4.20$, $MSe=.0133$, $p<.05$. Accuracy was 59.7% in the appropriate scene condition and 57.3% in the inappropriate scene condition. There was a main effect of target object manipulation, $F(3,69)=246.01$, $MSe=.0269$, $p<.05$. Accuracy was 80.3%, 26.4% and 47.7% in the same, deletion and orientation change conditions, respectively. Again, there was an interaction between "goodness of scene" and scene context, $F(1,22)=.02$, $MSe=.02$, $p>.05$. Responses were 59.6% and 57.9% correct in the appropriate and inappropriate scenes respectively for the 12 "best" scenes and 60.4% and 59.2% correct for the same conditions in the 12 remaining scenes. Also, there was an interaction between "goodness of scene" and target object manipulation, $F(3,69)=6.61$, $MSe=.0151$, $p<.05$. Finally, there was no 3-way interaction of goodness of scene, scene context, and target object manipulation, F's <1.

Insert Table 9 here

In the d' data, in the 12 best scenes, there was no main effect of scene context, $F(1,23)=.302$, $MSe=.542$, $p>.05$. There was, however, a main effect of target object manipulation, $F(1,23)=21.278$, $MSe=.203$, $p<.05$. Finally, there was a marginal interaction of scene context and target object manipulation, $F(1,23)=3.312$, $MSe=.123$, $p=.08$. In the 12 remaining scenes, there no main effect of scene context, $F(1,23)=1.313$, $MSe=.325$, $p>.05$. There was a main effect of target object manipulation, $F(1,23)=92.081$, $MSe=.216$, $p<.05$. And finally, no interaction, $F(1,23)=.160$, $MSe=.188$, $p>.05$.

In Experiment 4, the additional mean percentage correct and d' ANOVAs done with goodness of scene included in the analysis revealed the following (see Table 10). Like the analysis of Experiment 3's data, there was a main effect of goodness of scene, $F(1,23)=13.791$, $MSe=.0083$, $p<.05$. Participants' accuracy was 57.9% in the 12 "best" scenes and 61.3% in the remaining scenes. There was no main effect of scene context, $F(1,23)=.806$, $MSe=.0123$, $p>.05$. There was, however, a main effect of target object manipulation, $F(3,69)=426.94$, $MSe=.0277$, $p<.05$. Again, there was no interaction between "goodness of scene" and scene context, $F(1,23)=.308$, $MSe=.0108$, $p>.05$. There was an interaction between "goodness of scene" and target object manipulation, $F(3,69)=8.289$, $MSe=.0134$, $p<.05$. Also, there was an interaction of scene context and target object manipulation, $F(1,23)=2.180$, $MSe=.0097$, $p<.05$. Finally, there was no 3-

way interaction of goodness of scene, scene context, and target object manipulation, F's <1.

For the 12 best scenes, the d' data showed no main effect of scene context, $F(1,23)=.460$, $MSe=.417$, $p>.05$. There was a main effect of target object manipulation, $F(1,23)=75.79$, $MSe=.307$, $p<.05$. Also, there was a significant interaction, $F(1,23)=8.353$, $MSe=.145$, $p<.05$. In the 12 remaining scenes, there was no main effect of scene context, $F(1,23)=.035$, $MSe=.574$, $p>.05$. There was the main effect of target object manipulation, however, $F(1,23)=135.28$, $MSe=.240$, $p<.05$. There was, though, no interaction, $F(1,23)=2.551$, $MSe=.156$, $p>.05$.

Insert Table 10 here

An examination of the first three experiments suggests a significant context effect in the orientation change condition. To test for this effect, a subsidiary analysis was conducted on the orientation change condition treating experiment as a between-participants factor. In the overall analysis, the effect of context was highly reliable, $F(1,2)=10.678$, $MSe=.0015$, $p<.05$. Paired comparison, however, showed that the context effect in Experiments 2 and 3 was marginal, $F(1,23)=3.062$, $MSe=.1503$, $p=09$, and $F(1,23)=3.433$, $MSe=.1363$, $p=.07$ in Experiments 2 and 3, respectively, while the context effect was significant in Experiment 4, $F(1,23)=4.758$, $MSe=.0791$, $p<.05$.

Experiment 1 showed that participants did agree with the context manipulation, indicating that the scene context manipulation was effective. Target objects that I believed were appropriate for a given scene were indicated as appropriate by the participants. Target objects that I believed were inappropriate for a given scene, participants also believed were inappropriate. However, unlike the overall analyses conducted earlier, when the strength of the context manipulation for a given scene was factored into the analysis, the pattern of performance significantly changed in the same/difference task for participants' mean percentage correct and d' data. As a result, the data from the three experiments indicate that scene context does guide perceptual encoding, at least during a same/different decision task, though it can also influence post-perceptual encoding processes.

Experiment 5

Finding significant effects of scene context on perceptual encoding in the 12 best scenes offers support for a perceptual schema hypothesis as outlined above. Moreover, finding these effects with the same/different decision paradigm offers a greater level of certainty that the effects are not reflecting post-perceptual encoding processes. Previous research, while criticized for the conclusions drawn from them about the effects of scene context on object encoding, has found robust effects of context on eye-movements during scene viewing (Friedman, 1979; Boyce et al. 1989; Loftus and Mackworth, 1978). Therefore, in Experiment 5, participants viewed the scenes that had been used in Experiments 1-4 while their eye-movements were monitored. The purpose of this experiment was to determine if the pattern of eye-movements around the target object would be influenced by scene context, and more importantly, would this type of effect show up when a true measure of first fixation duration was used.

Method

Participants. Ten students at Michigan State University served as participants in this experiment. All participants received credit for participating in this experiment which served as partial fulfillment of their course requirements. Participants had normal vision, and had not participated in Experiments 1-4.

Apparatus. The stimuli were displayed at a resolution of 800 by 600 pixels on NEC Multisync XE 15" monitor driven by a Hercules Dynamite Pro super videographics adapter (SVGA) card. The screen refresh rate was 100 Hz. The contours of the objects and placeholders appeared black (pixels off) against a white (pixels on) background.

Eye-movements were monitored using a Generation 5.5 Stanford Research Institute Dual Purkinje Image Eyetracker (Clark, 1975; Cornsweet & Crane, 1973) which has a resolution of about 1' of arc and a linear output over the range of the visual display used. A bite-bar and forehead rest were used to maintain the participant's viewing position and distance. The position of the right eye was tracked, though viewing was binocular. Signals were sampled from the eyetracker by the computer using the polling mode of the Data Translations DT2802 analog-to-digital converter. This method of polling produced a sampling rate of better than 1 sample per millisecond.

Button-presses to begin the experimental trials were collected using a button panel connected to a dedicated input-output (I/O) card; depressing a button started a millisecond clock on the I/O card and generated a system interrupt that was serviced by software. The eyetracker, display monitor, and I/O card were interfaced with a microcomputer running a 66 MHZ 486 DX2 processor. The computer controlled the experiment and maintained a complete eye movement and button press record for each trial.

Materials. The 24 appropriate and the 24 inappropriate scenes used in Experiments 1-4 comprised the 48 scenes used in this experiment. None of the target object manipulation scenes were used.

Procedure. Participants were tested individually. Upon arrival at the experimental session, a bite bar for that participant was constructed. Once the bite bar had been constructed, the participant was seated in front of the computer monitor, eye-tracker and bite bar apparatus. They were then told that the purpose of this experiment was to examine how people look at scenes that they will have to later recognize. They were told that during the recognition test, they would have to distinguish between the original scenes and new scenes in which, for example, only a small detail of a particular object may have been changed. Participants were informed that on a given trial, the experimenter would press the button to start the trial and a scene would be presented to them for 15 seconds while their eye-movements were being monitored. Next, the experimenter would make sure the participant was still calibrated, and then press the button for the next trial. Before participants began the experimental trials, they would be calibrated on the eye-tracker and run in a set of practice trials. The calibration consisted of having the participant fixate 4 calibration markers at the top, bottom, left, and right sides of the display area. Calibration was checked by displaying a calibration screen consisting of six test positions and a fixation marker that indicated the computer's estimate of the current fixation position. The participant fixated the test positions, and if the fixation marker was +/-5 min arc of each, calibration was considered accurate.

Design and Analysis. Of a given scene, a participant saw either the appropriate or inappropriate scene condition during the experimental session. A given participant only saw 24 of the 48 possible scenes in this experiment, twelve in the appropriate context condition and twelve in the inappropriate context condition. Across participants, each scene appeared in each context condition an equal number of times. Participants' mean first fixation duration, total time, and gaze duration on the target object as well as percent entered and gaze duration counts were analyzed. First fixation duration was defined as the amount of time spent during the initial fixation on an object region and therefore excluded both intra-region and inter-region refixations. Gaze duration was defined as the sum of all fixation durations between first entry and first exit on an object region. Gaze fixation count was defined as the number of individual fixations between first entry and first exit for that region. Total fixation time was defined as the total amount of time spent fixating each object region during scene viewing. Total fixation count was defined as the total number of discrete fixations in the object region. Target object location regions were defined by constructing a box around the target object that was large enough to encompass both the appropriate and inappropriate object for a given scene. The pixel coordinates of the box were then used in the analysis program. The same box was used for both the appropriate and inappropriate context conditions for a given scene, so that the size of the scoring regions was equated across context conditions.

Eye Movement Data Analysis. Raw data files consisted of time and position values for each eyetracker sample. Because the analyses of interest are concerned with fixations, the saccades were removed from the data. Saccades were defined as velocities greater than 6.58 degrees per second. Manual inspection of the raw data files confirmed that this criterion was more effective at eliminating the initial and final stages of a saccade than were criteria of greater velocity. Once saccades had been eliminated, fixation positions and durations were computed over the remaining data. Fixation positions and durations were initially computed independently of the positions of the objects. The duration of a fixation was the elapsed time between two consecutive saccades. During a fixation, the eyes often drift. The position for a given fixation was taken to be the mean of the position samples (in pixel values) taken during that fixation weighted by the durations of each of those position samples, as given by the following equations:

$$xpos_{fix} = \frac{\sum (xpos_{sample} \times duration_{sample})}{\sum duration_{sample}}$$

$$ypos_{fix} = \frac{\sum (ypos_{sample} \times duration_{sample})}{\sum duration_{sample}}$$

Each fixation was then assigned to an object based on this position value.

Results. Figure 2 shows a typical scan pattern over a scene. Mean first fixation, gaze, and total time durations as well as percent entered fixation and gaze duration count for Experiment 5 are listed in Table 11. An ANOVA was conducted on each of these means, and in the interest of brevity, F-ratios for participants will be referred to as F1 and for items, F2.

Insert Figure 2 here

Participants entered the region containing the target object 95.8 % of the time in the appropriate scene context condition and 93.3% of the time in the inappropriate scene context condition. However, the main effect of context on the percentage of time participants entered the region containing the target object by scene context was not significant by participants nor by items, $F1(1,9)=.45$, $MSe=.007$, $p>.05$; and $F2(1,23)=1.30$, $MSe=.006$, $p>.05$. First fixation duration on the target object was not significantly different in the two scene contexts by participants and marginally significant by items, $F1(1,9)=2.18$, $MSe=2128.29$, $p>.05$ and $F2(1,23)=2.68$, $MSe=3031.54$, $p=.07$, respectively. Participants' mean first fixation durations were 296 msec in the appropriate scene context condition and 326 in the inappropriate scene context conditions.

Insert Table 11 here

Additional analyses were conducted on participants' first fixation duration data to determine when during viewing of the scene the target region was fixated for the first time. The number of fixations before initial fixation of the target object region was subjected to a median split and a tertiary split analysis. In the median split analysis, the number of fixations before the initial fixation of the target region was grouped into two sets: one through seven, and eight or more. In this analysis, there was no main effect of grouping or scene context, $F(1,9)=2.79$, $MSe=786.2778$, $p>.05$, and $F(1,9)=.7052$, $MSe=2859.2780$, $p>.05$, respectively. There was also no interaction, $F(1,9)=.3678$, $MSe=623.7500$, $p>.05$. These results suggest that the number of fixations on the scene before the first fixation of the target region has no effect on first fixation duration.

In the tertiary analysis, the fixations were grouped into three sets: one through four, five through ten, and 11 or more fixations. In this analysis, there was a marginal effect of grouping, $F(2,18)=3.1407$, $MSe=5063.6110$, $p=.07$. First fixation durations were 280.00, 330.25, and 282.95 msec for the first, second and third group, respectively. There was no main effect of scene context, $F(1,9)=.2352$, $MSe=8578.1670$, $p>.05$. However, there was a marginal interaction of grouping and scene context, $F(2,18)=3.0723$, $MSe=5166.0000$, $p=.07$. In the first group, first fixation duration was 305.10 and 254.90 msec in the appropriate and inappropriate scene context, respectively. In the second group, they were 300.20 and 360.300 msec for the appropriate and inappropriate context, respectively, and in the third group, they were 270.50 and 292.40 msec for the appropriate and inappropriate contexts, respectively. These results do suggest that the number of fixations on the scene before the first fixation of the target region does influence first fixation duration. Specifically, when the target region is quickly fixated during viewing, fixation duration is longer in the appropriate context than in the inappropriate context, but when the target region is fixated later during scene viewing, fixation duration is longer in the inappropriate context than in the appropriate context.

Participants' gaze durations differed significantly between the two scene context conditions for both participants, $F1(1,9)=8.48$, $MSe=31122.50$, $p<.05$, and items, $F2(1,23)=11.02$, $MSe=39665$, $p<.05$. Mean gaze durations on the target object were 472 and 702 msec in the appropriate and inappropriate scene context conditions, respectively. Gaze duration count also differed significantly between the two scene contexts, $F1(1,9)=6.71$, $MSe=.20$, $p<.05$ and $F2(1,23)=8.67$, $MSe=.24$, $p<.05$. The number of gaze fixations were 1.68 and 2.20 times for the appropriate and inappropriate scene context conditions, respectively.

The total time participants fixated the target object differed significantly between the two scene contexts, $F1(1,9)=12.38$, $MSe=291548.40$, $p<.05$; $F2(1,23)=19.74$, $MSe=417275.10$, $p<.05$. When the target object fit and did not fit the scene, total fixation times were 1308 and 2136 msec, respectively. Finally, fixation count differed significantly between the two scene contexts, $F1(1,9)=12.46$, $MSe=2.77$, $p<.05$ and $F2(1,23)=21.04$, $MSe=3.59$, $p<.05$. Participants fixated the target object 4.6 and 7.2 for the appropriate and inappropriate scene context conditions, respectively.

Discussion. The purpose of Experiment 5 was to test the manipulation of scene context in the scenes used in these experiments by replicating earlier findings from eye-movement studies. As stated earlier, research has found that eye-fixation patterns differ when a target object fits the scene than when it does not fit the scene, and the eye-fixation data reported here support this finding. However, as De Graef et al. (1990) suggest, this scene context effect is dependent on the type of measure used. Although calling it first fixation duration, most of the early eye-movement studies used a gaze duration measure

(Henderson, 1992a). The true measure of first fixation duration is the duration of time from the initial landing of the eyes on the target object *until another eye-movement is made*, including eye-movements made to another location on the target object (Henderson, 1992a). De Graef and his colleagues found that when true first fixation duration is used as a measure of encoding processes, scene context information does not influence eye-fixation data until around the 10th eye-fixation on the scene. In the present experiment, when participants were given 15 seconds to examine the scene, there was no reliable effect of scene context on first fixation durations, indicating that scene context did not affect early object encoding, though the pattern of data was in the correct direction. As De Graef et al. argue, fixating the target object region later during scene viewing leads to more of an effect of scene context on encoding. As a result, while there is a hint that first fixation durations are shorter in the appropriate scene context condition than in the inappropriate scene context condition, the lack of a reliable effect of scene context on first fixation duration could be because of a mixture of earlier and later viewing of the target object. However, examination of the number of fixations before the first fixation on the target region showed that when the target region was fixated did not reliably affect first fixation duration.

On the other hand, when other measures of eye-fixation are used, scene context information does influence eye-fixation patterns. For example, using gaze duration as a measure, research by Friedman (1979) and Loftus and Mackworth (1978) have found that fixations on a target object differ as a function of the probability of the object appearing in the scene. The same result was found here in the gaze duration, gaze duration count, fixation count and total fixation time data.

Thus, the results from Experiment 5 indicate that the scene context manipulation was sufficient for the purposes of the same/different task used in this set of experiments. The size of the context effect on gaze durations was of a similar magnitude to that found in these other studies. Moreover, when eye-fixation patterns are examined, an effect of scene context on object encoding is not found when more conservative measures are used (true first fixation) and is found when more general measures are used (gaze duration, for example).

General Discussion

Various studies have shown that scene context has some type of influence on object encoding, yet there is still a lack of a consensus on the nature of these effects. One widely held belief is that scene context guides perceptual object encoding, or those processes that take place up until the visual stimulus has been matched against its stored memory representation (Biederman, 1981; Loftus and Mackworth, 1978, Friedman, 1979; Boyce, Pollatsek, and Rayner, 1989; Boyce and Pollatsek, 1992), and this view has been summarized in the perceptual schema hypothesis. Unfortunately, prior research used to support the schema hypothesis has been criticized for possibly reflecting later post-perceptual processes such as the construction of a memory representation or checking to determine if the object makes sense in the scene. In fact, other research has shown that scene context does not guide object encoding but influences post-perceptual encoding (De Graef, Christiaens, and d'Ydewalle, 1990).

The purpose of this set of experiments was to address the question of scene context effects on perceptual encoding using a new paradigm. The same/different decision paradigm was chosen because it circumvented some of the concerns leveled against the prior research. Specifically, in this paradigm, it is less likely the participant can use guessing strategies, there is no possibility of priming from the name of the target object, and no need for the use of eye-movement recording or the naming of the target object by the participant. Here, participants were presented with a study scene, followed by a mask, and then a test scene with one object (the target object) changed or not. Their task was to determine whether the study and the test scenes were identical. When the two scenes differed, they differed in that the target object had undergone one of two possible manipulations: a deletion or an orientation change. The advantage of the same/different decision task was that there was no decision component in terms of identifying a target object, its location, or naming the target object. If the participants noticed the change, they must have encoded the information about the object. If appropriate context can enhance encoding, then they should notice the changes more when the object is consistent with the scene than when it is not.

In Experiment 1, the scene context manipulation used in the new paradigm was validated. Participants were shown the 24 appropriate and inappropriate scene context conditions (without seeing any of the scenes in which the target object had been manipulated) and were asked to determine if all of the objects "fit" the scene or not. Results showed that when the target object was in the scene considered inappropriate, participants judged them to be so. Likewise, when the target object appeared in the scene considered to be appropriate, participants judged them as so. These findings suggested that the scene context manipulation was sufficient for examining the effects of scene context on object encoding.

Scene Context Effects on Perceptual Encoding: Evidence from the new paradigm.

In Experiment 2, the effects of scene context on perceptual encoding were examined. In this experiment, an initial scene presentation time of 250 msec was used in an attempt to address perceptual encoding of information in a scene. There was no significant effect of scene context on perceptual encoding in the percent accuracy data. However, there was a marginal main effect of scene context in the d' data. Participants were slightly more sensitive to the target object manipulations in the appropriate scene context condition than in the inappropriate scene context condition. However, when the strength of the scene context manipulation for a given scene was factored into the analyses, there were effects of scene context on encoding of the target object for both the percent correct and the d' data. As such, these results offer support for a perceptual schema hypothesis. An interesting finding from Experiment 2 was the difficulty participants had in detecting the deletion manipulation in both the appropriate scene context and the inappropriate scene context. Further discussion of this finding will follow.

Previous studies using stimuli very similar to those used here have demonstrated that the "gist" or meaning of a scene can be accessed within the first 150 ms of scene viewing (Biederman, 1981; Biederman et al. 1982; Boyce et al, 1989). While, the results of Experiment 2 suggest that scene context exerts an influence on initial perceptual encoding, Experiment 3 examined the effects of scene context on object processing using a study scene presentation time of 500 msec to determine if a larger scene context effect would result. In Experiment 3, there was no effect of scene context and no interaction of context and the target object manipulations in either the percent accuracy and the d' data. However, when strength of the scene context manipulation was factored into the analysis, there was a main effect of scene context in the percent correct data but not in the d' data. As with Experiment 2, participants detected the orientation change much more than they did the deletion. Therefore, it appears to be the case that scene context does influence early perceptual encoding as predicted by the perceptual schema hypothesis.

While scene context appears to influence perceptual encoding, Experiments 2 and 3 did not examine this effect on post-perceptual processing.. To examine this question, in Experiment 4, the study scene presentation time was increased from 500 msec to 2500 msec. With a presentation time of this length, it was most likely that perceptual encoding was no longer being addressed, but post-perceptual encoding, such as the time needed to create a memory representation of the scene would be. In De Graef's earlier work, he found that scene context information does have an effect on object encoding at around the tenth eye-fixation.

In Experiment 4, while there was neither a significant main effect of scene context nor an interaction of scene context and target object manipulation in the percent correct data, there was a significant interaction of scene context and target object manipulation in the d' data. In this experiment, participants detected the orientation change manipulation better in the appropriate scene context condition than in the inappropriate scene context condition. However, like the two prior experiments, when strength of the scene context manipulation was factored into the analyses, the pattern of results changed. There was a significant interaction of scene context and target object manipulation in the d' data for the 12 best scenes. In the percent correct data, there was a marginal effect of scene context and target object manipulation.

Finally, in Experiment 5, the eye-movement data indicated that scene context did not reliably influence early object encoding, as measured by first fixation duration. However, it is not known when participants were fixating the object, i.e., early during scene viewing or later during the scene viewing. As a result, while there is a hint that first fixation durations are shorter in the appropriate scene context condition than in the inappropriate scene context condition, the lack of a reliable effect of scene context on first fixation duration could be because of a mixture of earlier and later viewing of the target object. However, a regression analysis showed that the number of fixations before the target region was fixated did not influence first fixation duration. Moreover, grouping the number of fixations before fixation of the target region did not produce reliable effects on first fixation duration.

Detecting Orientation Changes Vs. Deletions

An interesting finding from Experiments 2, 3 and 4 is the lack of detection of the deletion manipulation on the part of the participants. In fact, across the three experiments, the detection of the deletion did not vary significantly while detection of the orientation change manipulation increased (see Figure 3, which shows the d' data and 95% confidence intervals for the detection of the deletions and orientation changes as a function of experiment). At the outset, this finding seems highly counterintuitive. When an object is deleted from a scene, a change occurs in all of the types of visual information outlined by Mandler and Johnson present in a scene. Consequently, it would seem likely that with such a disruption of information, this type of difference in the scene would be readily apparent. However, as is the case with the present studies, participants often fail to notice when some detail in a scene has been deleted from the scene (Hearst, 1991; Agostinelli, Sherman, Fazio, & Hearst, 1986; Pezdek, Maki, Valencia-Laver, Whetstone, Stoeckert, & Dougherty, 1988). For example, Pezdek et al. (1988) examined participants' recognition memory for pictures, assessing memory for the addition or deletion of specific details in the pictures. In their study, participants were given a sentence prompt or no sentence prompt and then presented with either simple or complex line drawings of pictures and later given a same/changed recognition memory test. Both the simple and the complex version of a given picture could be described by the same sentence. For the addition condition, extra shading, details, and elaboration were added to the simple version of the picture. In the deletion condition, the extra shading, etc. was deleted from the complex version of the picture. Participants were presented with either the same picture at study and test, or with the simple version followed by the complex (addition condition) or the complex followed by the simple (deletion condition). Pezdek et al. posited that the sentence prompt would increase the likelihood that the pictures would be processed in terms of their central schema. They found what they referred to as the *asymmetric confusability effect* (Pezdek and Chen, 1982, cited in Pezdek et al. 1988) or the finding that participants' d' values in detecting the changes were greater for additions than for deletion conditions. Moreover, they found that the sentence prompt condition exaggerated this effect. According to this effect, during the study phase, pictures are encoded such that both complex and simple versions are represented in memory as the simple version. Thus, deleted detail in the test scene is difficult to detect because the complex version containing the detail was encoded like the simple version in the memory representation. In the case of additions, the simple version is encoded during the study phase and differs from the test scene with the added detail, thus easier to detect.

Insert Figure 3 here

What is interesting about the findings from these earlier experiments and the results from the experiments reported here is that in the present experiments, participants did not notice deletions in either the appropriate context or the inappropriate context conditions. Friedman (1979) reported that participants notice changes to nonobligatory objects more than they do the same changes to obligatory objects in a scene, a result that is also different from the Pezdek et al. (1982) findings. Pezdek posits that differences in the magnitude of the schemata manipulation could account for the differences between their results and Friedman's. This possibility could also explain the present findings even though Experiment 1 indicated that participants were aware of the difference between the appropriate and inappropriate scene context conditions, suggesting that given the chance, they should have detected the deletion in the inappropriate scene context condition.

But why is detection of the orientation change manipulation better than detection of the deletion manipulation in these experiments? One possibility has to do with the presence or absence of a retrieval cue. In the orientation change condition, when the test scene is presented, the presence of the target object (albeit slightly changed) serves as a retrieval cue for the same object in the preceding study scene. In this case, the same/different decision can proceed based on the memory representation of the first scene and the perceptual representation of the second (test) scene. Performance in detecting the orientation change increases with increased study scene presentation time (as can be seen in Figure 3) because of the increase in the amount of time available to construct a memory representation of the study scene. And it is at the longest study scene presentation time that scene context information has an effect on detection of the orientation change manipulation. Context can have an effect on detection of the orientation change at the longest display duration because context has exerted an effect on the memory representation of the object, and this memory representation is being retrieved by the object's presence in the second display

In the deletion manipulation, when the test scene is presented, there is no target object to serve as a retrieval cue for the object in the study scene. As a result, the deletion manipulation is not detected. While it is possible that the empty space or new contours created by the deletion of the target object in the test scene could serve as a weak retrieval cue, the data do not bear this out. Because there is no cue to access the memory representation of the object in the deletion condition, it follows that there would be no effect of scene context. This is because, even if scene context did have an influence on that representation, it would not be manifested because the representation does not get accessed. In other words, at the long display duration, context influences memory encoding, but this only shows up in the orientation change condition because it is the only condition that actually taps into (retrieves) the memory representation.

The Effects of Scene Context on Encoding an Object's Orientation.

As stated above, in Experiment 2,3 and 4, the orientation change manipulation was better detected in the appropriate scene condition than in the inappropriate scene condition. Concerning this effect of context on the orientation change manipulation, proponents of the perceptual schema hypothesis differ. While the perceptual schema hypothesis predicts that the gist of a scene will be apprehended quickly and guide perceptual encoding, it does not specify the direction of the influence. In other words, will an appropriate scene context make it easier or more difficult to perceptually encode information about an object? Participants' detecting the orientation change better in the appropriate scene context is not consistent with Friedman's frame theory (1979). According to the frame theory, this type of target object manipulation should be better detected when the target object was in the inappropriate scene context condition. According to the frame theory, when an object does not fit the scene, more effortful processing of the object occurs. This additional processing allows for more specific information about the object to be encoded, including, for example, the direction that the object is facing. Consequently, detection of a change in some specific feature about the object should be more easily detected in an inappropriate scene context.

The present scene context and target object interaction can be explained by Biederman and his colleagues' view of the perceptual schema hypothesis. According to Biederman and his colleagues (1981, 1982), scene context facilitates encoding of information about an object that belongs in that scene. In this case, one can assume that when an object does not fit a scene, only partial perceptual encoding of the object can result quickly. This fact may be because the observer is trying to figure out what the object is and/or how it fits into the scene and does not begin to encode specific information about the target object, such as its orientation, until later. (For example, that's odd that there was a bicycle in the grocery store, but I can't remember what direction it was facing). Or, possibly, when the object does not fit, information about the object is not encoded at all, although this possibility seems unlikely in this experiment in that participants were detecting the orientation change manipulation in the inappropriate scene context condition above chance. Nevertheless, according to this view, when the object does fit the scene, specific information about the object is readily encoded, so that changes to specific information about the object is more easily detected. This type of explanation fits with the finding that the orientation change was better detected when the target object appeared in the appropriate scene. In this case, when the target object is in the appropriate scene context, encoding of information, including its orientation, is facilitated.

So in conclusion, a perceptual schema hypothesis as argued by researchers like Biederman et al., (1981, 1982) and Boyce et al., (1989, 1991) can account for the data reported from these experiments. Finding a reliable effect of scene context at 250 msec can only be explained by a schema hypothesis that posits that scene context influences the encoding of perceptual processing. Moreover, finding these effects using the same/different paradigm offers converging evidence that scene context influences perceptual encoding, evidence from a paradigm that circumvents some of the problems leveled against prior research. As such, these results support the hypothesis that scene context influences perceptual encoding processing as well as some post-perceptual processing.

Table 1.

Omnibus ANOVA: Participant's Mean Percentage Correct Responses for Experiment 2.

(Mean d' in parenthesis)

Scene Context	Target Object Manipulation			
	Same	Deletion	Orientation change	Mean
Appropriate	77.3	31.6 (.30)	53.8 (.92)	60.0
Inappropriate	74.5	30.8 (.19)	49.7 (.73)	57.4
Mean	75.9	31.2	51.7	

Table 2.

Quartile Analysis of Mean Percentage Correct Responses for Experiment 2.

1st Quarter of Experimental Trials.

Scene Context	Target Object Manipulation			
	Same	Deletion	Orientation change	Mean
Appropriate	73.6	41.3	60.9	58.6
Inappropriate	70.7	36.4	50.9	52.7
Mean	72.2	38.9	55.9	

4th Quarter of Experimental Trials.

Scene Context	Target Object Manipulation			
	Same	Deletion	Orientation change	Mean
Appropriate	78.0	24.9	48.6	50.5
Inappropriate	77.5	35.3	49.4	54.1
Mean	77.8	30.1	49.0	

Table 3.

Omnibus ANOVA: Participant's Mean Percentage Correct Responses for Experiment 3.

(Mean d' in parentheses)

Scene Context	Target Object Manipulation			
	Same	Deletion	Orientation change	Mean
Appropriate	80.4	27.4 (.32)	50.7 (.96)	59.7
Inappropriate	79.5	25.5 (.22)	43.7 (.76)	57.1
Mean	80.0	26.5	47.2	

Table 4.

Quartile Analysis of Mean Percentage Correct Responses for Experiment 3.

1st Quarter of Experimental Trials

Scene Context	Target Object Manipulation			
	Same	Deletion	Orientation change	Mean
Appropriate	79.1	35.0	52.3	55.5
Inappropriate	82.4	21.8	52.9	52.4
Mean	80.7	28.4	52.6	

4th Quarter of Experimental Trials

Scene Context	Target Object Manipulation			
	Same	Deletion	Orientation change	Mean
Appropriate	85.7	18.3	52.2	52.1
Inappropriate	85.8	19.3	44.8	50.0
Mean	85.8	18.8	48.0	

Table 5.

Omnibus ANOVA: Participant's Mean Percentage Correct Responses for Experiment 4.

(Mean d' in parentheses)

Scene Context	Target Object Manipulation			
	Same	Deletion	Orientation change	Mean
Appropriate	88.1	14.6 (.14)	49.5 (1.26)	60.1
Inappropriate	87.5	15.3 (.18)	44.8 (1.12)	58.8
Mean	87.8	15.0	47.1	

Table 6.

Quartile Analysis of Mean Percentage Correct Responses for Experiment 4.

1st Quarter of Experimental Trials

Scene Context	Target Object Manipulation			
	Same	Deletion	Orientation change	Mean
Appropriate	86.9	21.7	66.9	71.3
Inappropriate	84.2	17.6	62.3	72.1
Mean	85.5	19.6	64.6	

4th Quarter of Experimental Trials

Scene Context	Target Object Manipulation			
	Same	Deletion	Orientation change	Mean
Appropriate	87.2	8.7	43.7	80.5
Inappropriate	86.9	10.0	38.1	81.4
Mean	87.0	9.4	40.9	

Table 7.

Median Split Ranking of the 24 scenes:

12 "Best" Scenes:	12 "Remaining" Scenes:
Bar Scene (.938)	Bathroom (.625)
Bedroom (.875)	Beach (.250)
Bus Station (.938)	Chemistry Lab (.562)
Checkout Counter (.813)	Classroom (.562)
Church (1.00)	Farm (.500)
Construction Site (1.00)	Kitchen (.625)
Dining Room (.938)	Laundry (.750)
Dock (.876)	Living Room (.813)
Gas Station (.876)	Office (.813)
Library (.876)	Pool (backyard) (.813)
Locker Room (1.00)	Restaurant (.687)
Theatre (1.00)	Workshop (.745)

Table 8.

Median Split Analysis of Mean Percentage Correct Responses for Experiment 2.

(D' in parentheses)

12 "Best" Scenes

Scene Context	Target Object Manipulation			
	Same	Deletion	Orientation change	Mean
Appropriate	77.6	33.3 (.39)	53.8 (1.02)	54.9
Inappropriate	73.3	31.5 (.17)	42.7 (.46)	49.2
Mean	75.4	32.4	48.2	

12 "Remaining" Scenes

Scene Context	Target Object Manipulation			
	Same	Deletion	Orientation change	Mean
Appropriate	76.9	29.9 (.20)	53.8 (.91)	53.5
Inappropriate	75.6	30.2 (.24)	56.6 (1.03)	54.1
Mean	76.3	30.1	55.2	

Table 9.

Median Split Analysis of Mean Percentage Correct Responses for Experiment 3.

(D' in parentheses)

12 "Best" Scenes

Scene Context	Target Object Manipulation			
	Same	Deletion	Orientation change	Mean
Appropriate	80.2	28.5 (.27)	46.9 (.83)	51.9
Inappropriate	80.4	26.9 (.32)	37.2 (.61)	48.2
Mean	80.3	27.7	42.1	

12 "Remaining" Scenes

Scene Context	Target Object Manipulation			
	Same	Deletion	Orientation change	Mean
Appropriate	80.6	26.4 (.21)	54.5 (1.15)	53.8
Inappropriate	78.7	26.9 (.11)	52.4 (.99)	51.7
Mean	79.7	25.2	53.5	

Table 10

Median Split Analysis of Mean Percentage Correct Responses for Experiment 4.

(D' in parentheses)

12 "Best" Scenes

Scene Context	Target Object Manipulation			
	Same	Deletion	Orientation change	Mean
Appropriate	88.6	12.5 (-.07)	45.1 (1.14)	48.7
Inappropriate	88.9	13.8 (.24)	36.8 (1.00)	46.5
Mean	86.8	13.1	41.0	

12 "Remaining" Scenes

Scene Context	Target Object Manipulation			
	Same	Deletion	Orientation change	Mean
Appropriate	87.6	16.7 (.17)	54.2 (1.46)	52.8
Inappropriate	86.3	19.1 (.27)	52.8 (1.30)	52.7
Mean	87.0	17.9	53.5	

Table 11.

Eye-Movement Measures Data from Experiment 5.

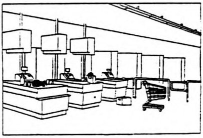
Eye-movement Measure	Appropriate Scene Context	Inappropriate Scene Context
Percent Entered	95.8%	93.3%
First Fixation	296 ms	326 ms
Gaze Duration	472 ms	702 ms
Number of Gaze Fixations	1.68	2.20
Total Time	1339 ms	2189 ms
Number of Fixations	4.60	7.23

Figure 1.

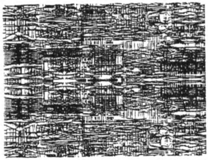
An example trial sequence for Experiments 2,3, and 4.

Fixation Cross (500 msec)

Study Scene (250, 500, or 2500 msec)



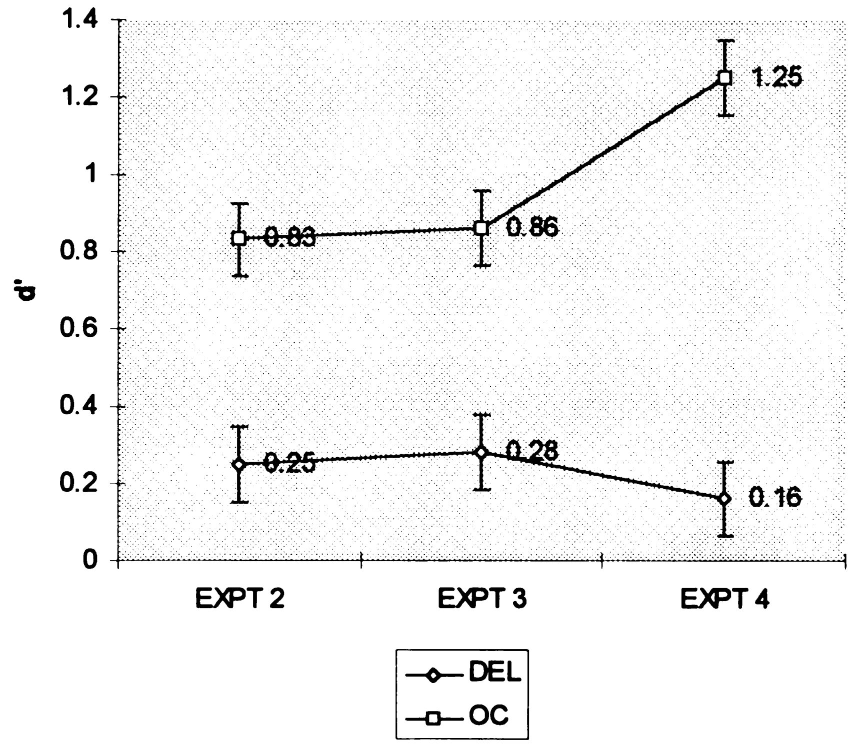
Mask (400 msec)



Test Scene (until participant responds)

Figure 3.

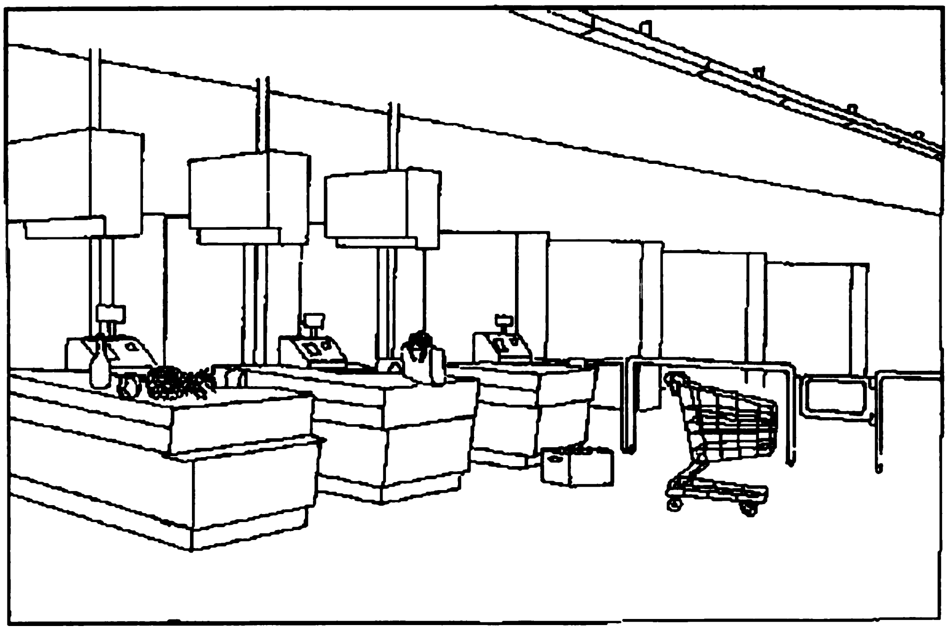
Mean D' for Target Object Manipulations Across Experiments 2, 3, and 4.



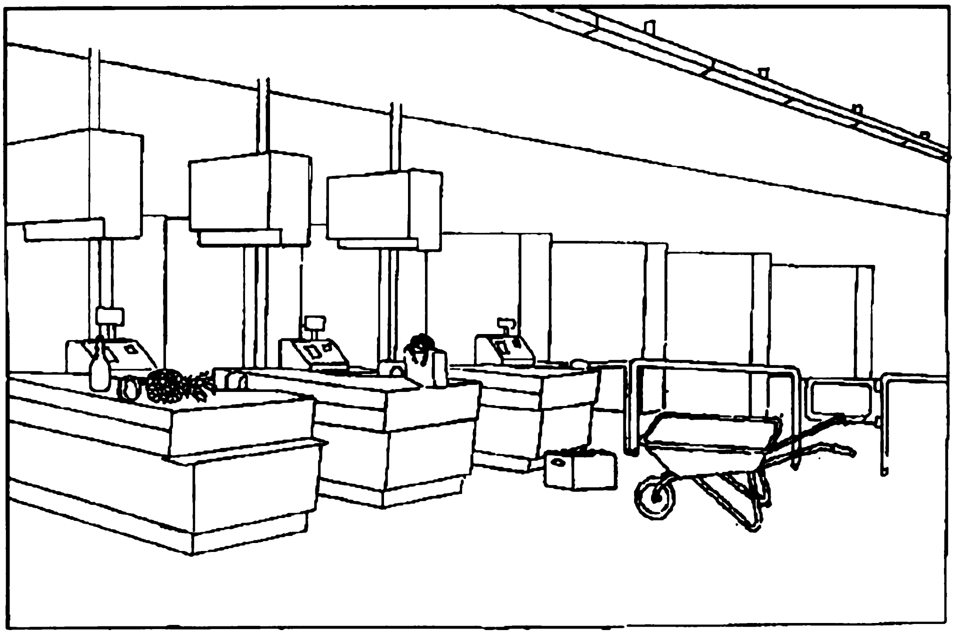
Appendix A

Appendix A contains two examples of the 48 appropriate scene contexts and inappropriate scene context scenes used in these experiments. The two scenes are the checkout counter scene, with the grocery cart as the appropriate target object and the wheel barrel as the inappropriate target object, and the backyard scene with the wheel barrel as the appropriate target object and the grocery cart as the inappropriate target object. The orientation change and the deletion conditions are not shown.

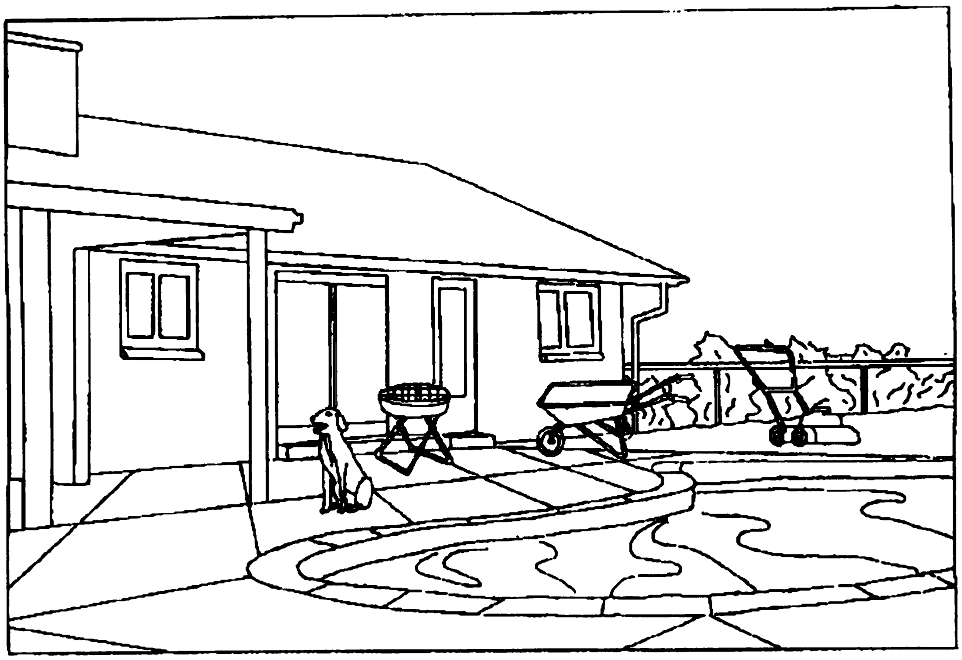
Checkout Counter Scene: Appropriate Scene Context



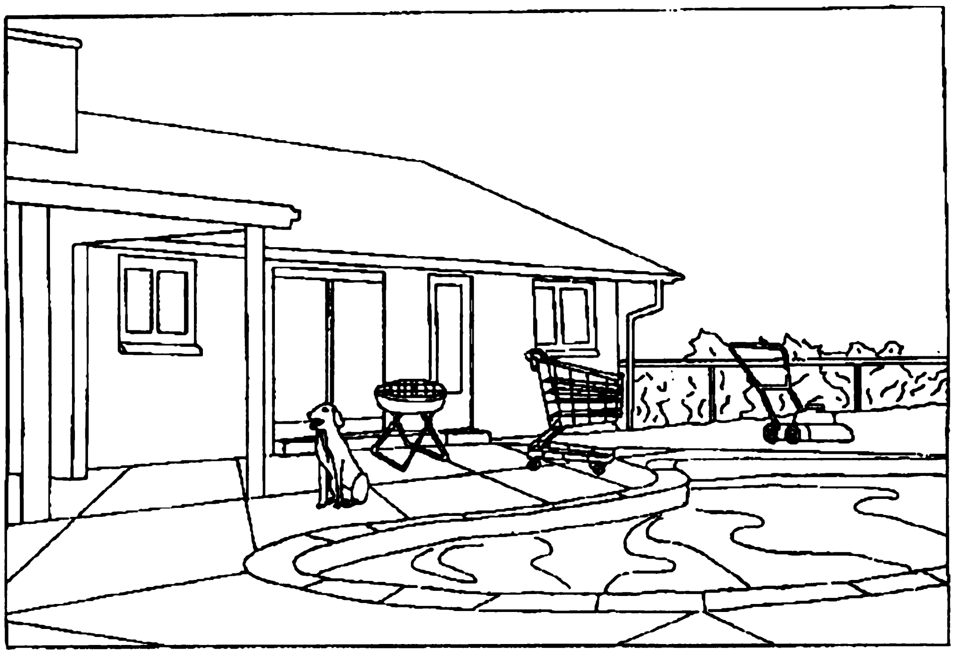
Checkout Counter Scene: Inappropriate Scene Context



Backyard Scene: Appropriate Scene Context



Backyard Scene: Inappropriate Scene Context



References

Agostinelli, G.; Sherman, S.J.; Fazio, R.H.; and Hearst, E.S. (1986). Detecting and identifying change: additions versus deletions. *Journal of Experimental Psychology: Human Perception and Performance, 12,* 445-454.

Antes, J. R. (1974). The time course of picture viewing. *Journal of Experimental Psychology, 103,* 62-70.

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

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

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

Biederman, I.; and Ju, J. (1988). Surface versus edge-based determinants of visual recognition. *Cognitive Psychology, 20,* 38-64.

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.

Boyce, S. J.; & Pollatsek, A. (1992). An exploration of the effects of scene context on object identification. In, Rayner, K.; (Ed), *Eye movements and Visual Cognition.* Springer-Verlagl New York.

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.

Coren, S.; Porac, C.; & Ward, L. M. (1984). *Sensation and Perception.* Harcourt Brace Jovanovich: Chicago.

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

De Graef, P. (1992). Scene-context effects and models of real-world perception. In K. Rayner, (Ed.). *Eye movements and Visual Cognition.* Springer-Verlagl New York.

De Graef, P.; DeTroy, A.; & d'Ydewalle, G. (1992). Local and global contextual constraints on the identification of objects in scenes. Special Issue: Object perception and scene analysis. *Canadian Journal of Psychology, 46,* 489-508.

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

Hearst, E. (1991). Psychology and nothing: recognizing and learning from absence, deletion, and nonoccurrence are surprisingly difficult. Animals and people, it seems, accentuate the positive. *American Scientist, 4,* 432-443.

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

Henderson, J. M., Pollatsek, A., & Rayner, K. (1989). Covert visual attention and extrafoveal preview on object identification. *Perception & Psychophysics, 45,* 196-208.

Henderson, J. M. (1992a). Object identification in context: The visual processing of natural scenes. *Canadian Journal of Psychology, 46:3,* 319-341.

Henderson J. (1992b). Visual attention and eye movement control during reading and picture viewing. In K. Rayner, (Ed.) *Eye movements and Visual Cognition.* Springer-Verlagl New York.

Intraub, H. (1980). Presentation rate and the representation of briefly glimpsed pictures in memory. *Journal of Experimental Psychology: Learning, Memory and Cognition, 15,* 179-187.

Intraub, H. (1984). Conceptual masking: the effects of subsequent visual events on memory for pictures. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 10,* 115-125.

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; 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.

Mackworth, N. H.; & Morandi, A. J. (1967). The gaze selects informative details within pictures. *Perception and Psychophysics, 2,* 547-552.

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*, 509-522.

Mandler, J. M. & Ritchey, G. H. (1977). Long-term memory for pictures. *Journal of Experimental Psychology: Human Learning and Memory, 3*, 386-396.

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

Pezdek, K.; Maki, R.; Valencia-Laver, D.; Whetstone, T.; Stoeckert, J.; & Dougherty, T. (1988). Picture memory: recognizing added and deleted details. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14,* 468-476.

Potter, M. C. (1976). Short-term conceptual memory for pictures. *Journal of Experimental Psychology: Human Learning and Memory, 81*, 10-15.

Sekuler, R. Blake, R. (1990). *Perception.* McGraw-Hill: New York.