EPISODIC PRIMING
AND
OBJECT PROBABILITY EFFECTS

A Thesis Presented

by

PETER DE GRAEF

Submitted to the Graduate School of the
University of Massachusetts in partial fulfillment
of the requirements for the degree of

MASTER OF SCIENCE

September 1990

Department of Psychology

EPISODIC PRIMING
AND
OBJECT PROBABILITY EFFECTS

A Thesis Presented

by

PETER DE GRAEF

Approved as to style and content by:

Keith Rayner

Keith Rayner, Chairperson of Committee

Alexander Pollatsek

Alexander Pollatsek, Member

Rachel Clifton

Rachel Clifton, Member

Melinda Novak

Melinda Novak, Department Head
Department of Psychology

ACKNOWLEDGEMENT

I would like to thank Keith Rayner for taking me on as one of his students and having given me the opportunity to expand my views and get a much firmer grip on cognitive psychology. In addition, I greatly appreciate his patience and continuing interest during the slow birth of the present manuscript. I also want to thank Rachel Clifton for serving on my committee and Sandy Pollatsek for his helpful comments and views about this and related aspects of my work. Finally, I wish to express my gratitude to all faculty and students of the Cognitive and Developmental area who have made my stay in Amherst a very stimulating experience and a source of pleasant memories.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Figure | Page

INTRODUCTION

Most prominent models of visual object perception (e.g. Biederman, 1987; Marr, 1982) view the apprehension of a particular object in an image as exclusively based on a data-driven and pre-conceptual recovery of the object's structural features (i.e. geons, contoursegments, etc.) from the image. Research on object perception in full scene context, however, has suggested that this view may need to be modified if one wishes to model more than the perception of unanticipated, isolated objects. The purpose of the present thesis is to contribute to an evaluation of the degree to which this challenge should be taken seriously.

The first chapter of this thesis will present and discuss the two main lines of research developed in the study of effects of scene context on object identification. An initial section in this chapter will deal with research concentrating on the question whether scene context has any influence on the nature of the pattern recognition processes involved in object identification. The central assumption in this research is that prior to or during the first glance at a real-world scene, a scene-specific schema or frame is activated which provides an integrated representation of the typical makeup and contents of the viewed scene. Based on this assumption, several authors have advanced the hypothesis that the identification of objects with a high probability of appearance in a given scene, is based on the concept-driven and resource-inexpensive detection of global object features specified in the frame representing that scene. This in contrast, they propose, to either improbable or isolated objects which are identified on the basis of a data-driven and resource-demanding analysis of visual detail. While the studies designed to test this hypothesis have generally yielded results which have been taken to corroborate its validity, a detailed examination of the evidence will be presented in order to demonstrate that this conclusion may be unwarranted.

A second section will review research which emphasizes that scene context affects object identification by providing a frame of reference in which spatial object-context relations define a set of relational object features which are used as a basis for object recognition. Based on an analysis of what distinguishes a well-formed, natural scene from an unstructured array of isolated objects, it has been argued that an object's appearance in such a scene can not only be characterized as probable, but also conforms to a limited set of fundamental spatial object-context relations. The concrete realization of these relations in the appearance of a particular object in a particular scene is considered to be quite stable across instances of that object-scene combination. Consequently, an object's typical spatial relations to a scene it is likely to appear in, are taken to be an integral part of the global schema for that scene. Under the assumption that such schemas are inevitably activated in the earliest stages of scene exploration, it has been hypothesized that relational object features are an integral part of the image information used for object identification in scenes. Since a number of studies examining this hypothesis appear to indicate that violations of spatial object-context relations decrease the identifiability of objects, its validity has generally been accepted. Again however, a detailed discussion of this research will be presented in order to demonstrate that this conclusion may not be justified.

Based on the review presented in the first chapter, it will be argued that only an object's probability of appearing in a scene can safely be regarded as having a genuine effect on the ease with which the object can be identified. One can therefore pose the question of how existing models of object perception should be modified in order to account for this contextual effect. In order to answer this question, it is necessary to decide between two alternative views that have been proposed in order to account for the object probability effect. On the first view, the effect reflects a top-down influence of a global scene-schema, resulting in the concept-driven identification of individual objects. If correct, this would necessitate a drastic revision of the presently accepted data-driven accounts of object perception. On the second view, however, such a revision would not be required since it attributes the object probability effect to the operation of a passive priming mechanism between the individual object representations that are used to categorize the object models computed through data-driven feature analysis of an image. The remainder of the thesis therefore will focus on evaluating the validity of this simpler and more conservative inter-object priming account.

Specifically, the second chapter will identify three possible constraints on this priming mechanism, which, if proven to be true limits, could serve as basis for an empirical and unequivocal test of the mechanism's validity as an explanation of context effects in full scenes.

Finally, in the third chapter two experiments are reported which were designed to test the existence of these three possible constraints. Based on the results obtained in these experiments, some conclusions are offered with regard to the nature of the inter-object priming mechanism and a test of its role in an account of real-world scene perception.

CHAPTER 1

EFFECTS OF SCENE CONTEXT ON OBJECT IDENTIFICATION: A REVIEW OF THE EVIDENCE AND THEORIES

1.1 Effects of Scene Context on Pattern Recognition Processes in Object Identification

As mentioned earlier, most research on scene context effects has been inspired by the assumption that real-world scene perception is mediated by scene-specific schemas, activated prior to or during the first few glances at a scene. The rationale behind this central assumption and the predictions that have been derived from it with respect to the context-sensitivity of object pattern recognition, have been outlined most clearly by Friedman (1979).

In her frame theory of scene perception, Friedman sets out from the position that apprehending natural scenes and their components from arrays of optical information requires an interaction between the output of low-level feature analyzers and a priori knowledge about how those features go together and what scene(component) these feature combinations signify. This a priori knowledge, she assumes, takes on the form of frames (or schemas) which each constitute a representation of a particular reality at a specific level of abstraction or globality (e.g. shape-frames, volume-frames, object-frames, place-frames, scene- frames, etc.). At the basis of this assumption lies the view that frames -as outlined in various theories of world knowledge representation (e.g. Bobrow & Norman, 1975; Minsky, 1975; Palmer, 1975; Schank & Abelson, 1977)- have certain properties which allow them to function as powerful pattern interpreters.

The first one is that they employ an abstract representational format (i.e. propositions or procedures) which allows for the integration into one frame of the viewer's knowledge about both the semantic and physical characteristics of its real-world referent. Consequently, as Palmer (1975) points out, no resources need to be spent in translating visual information into a code allowing for its meaningful interpretation.

A second property is that they represent a particular reality in a prototypical fashion, i.e. in terms of both its invariant characteristics and limited ranges or probabilistic distributions of concrete values its variable characteristics can take on. This implies that, given sufficiently broad experience with exemplars of scenes and their components, i) a limited number of these frames will provide sufficient power and flexibility to interpret a wide variety of feature patterns, and ii) these representations can serve as a basis for generating accurate expectations about the visual and semantic characteristics the instances of its referent are likely to have. Consequently, if it would be possible to access a frame prior to an extensive pattern analyis of a scene or object, then this could substantially reduce the time and effort required for the recognition of that scene or object. This because such access would provide a frame of reference for generating expectations that can constrain the universe of all possible pattern tests to the subset of those that are most likely to lead to a coherent interpretation of the pattern at hand.

According to Friedman, it is precisely a third general property of frames which enables this kind of access. Specifically, this property is that frames (as a consequence of the abstract representational format they employ - Fischler, 1978-) need to specify their referents in a relative fashion. For this purpose they draw upon a varied repertoire of physical and semantic relations (e.g. probability of co-occurence, relative size and location, part-structure, properties, class membership, etc.). This aspect of frames is illustrated in Figure 1 which represents (part of) what a 'face-frame' could look like.

The important thing to note about this frame representation of a face is that it not only stipulates overall face properties (i.c. shape but others like for instance color or dimensionality may be added). Indeed, it also makes explicit the face's *internal and external* structure by defining it as *having parts* (eyes, nose and mouth) which each have properties of their own as well as a particular size, location

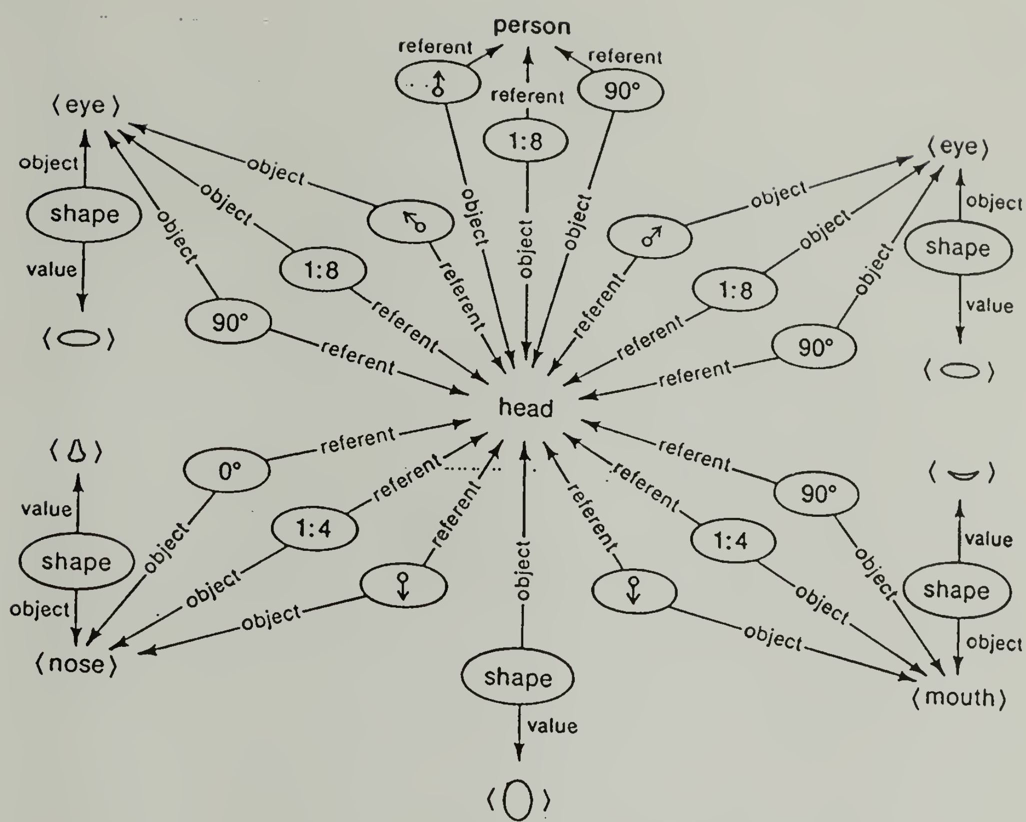

Figure 1. Representation of a 'face-frame' according to Palmer (1975).

-the vector symbols- and orientation relative to the face, and as *being a part* of a person with a particular size, location and orientation relative to that person. Furthermore, it should be noted that each argument in a frame can itself be considered to be the referent of another frame, which allows for the organization of frames into systems representing scenes and their components at multiple levels of abstraction or globality (Hanson & Riseman, 1978; McArthur, 1982; Palmer, 1975). With respect to the issue of frame access this implies that a particular frame can be accessed and constrain further processing either on the basis of a partial or lower-level analysis of a pattern instance of its referent, or on the basis of expectations generated by frames representing the visual and semantic context of its referent.

In order to examine the validity of this frame theory of scene perception, Friedman outlined and tested its implications for object identification in scenes. Specifically, she proposes that scenes are rarely encountered out of context and that consequently the appropriate scene-frame will generally be accessed prior to the actual viewing of the scene. Since this scene-frame specifies the scene's prototypical internal structure (i.e. objects and background components that typically appear in the scene as well as relations that usually hold between them) its activation will generate expectations about what objects are likely to be present in the scene and what the typical features of these objects are. Given the fact that the object representations activated in this manner are arguments in a global scene-frame rather than frames which are fully expanded at the object level, they will specify only those object features which need to be detected in addition to the available contextual information in order to establish the presence of the object in question in that scene. Specifically, Friedman contends that global object features (e.g.

shape, dimensionality, texture, etc.) are sufficient for this purpose. Consequently, she claims, the identification of an object in a scene it is likely to appear in, is generally based on a rapid and resource-inexpensive detection of its global features suggested by the previously activated frame for that scene. Alternatively, if an object is unlikely to appear in a given scene the prior activation of the corresponding scene-frame is to no avail and the object will have to be identified on the basis of a slower and more resource-demanding interaction between a more detailed feature analysis and a frame which is fully expanded at the object level of representation.

In order to test this hypothesis, Friedman presented subjects with two series of pictures. The first contained line drawings of scenes which each were shown for 30 sec after the subjects had been cued with the general theme of the scene in order to activate the hypothesized scene-frame. During the viewing of these scenes eye movements were recorded. The second series contained the same scenes along with distractor versions in which individual objects had been altered, relocated, substituted or deleted. Prior to the viewing of the first set, subjects were informed that they would have to be able to distinguish the scenes they were about to see from new scenes in which, for example, only a small object detail would be different. For all objects presented in the pictures, ratings of their a priori probability of appearance in the scene had been collected prior to the experiment.

With regard to the nature of object pattern recognition processes, Friedman predicted two things in this experiment. First, under the assumption that the duration of the first fixation on an object (FFD) reflects the time needed to encode and identify an object, she expected that the hypothesized difference between rapid global feature detection and time-consuming detailed feature analysis should be reflected in longer first fixations on improbable than on probable objects. Second, she predicted that this difference should also lead to superior memory for the details of improbable objects and that consequently distractors in which improbable objects had been slightly altered should be discriminated more accurately from the original scenes.

An analysis of the recorded eye movement patterns and recognition data confirmed both predictions. Apparently, this supports the theory that placing an object in a consistent scene does indeed alter the object's pattern recognition from a data-driven process of detailed feature analysis to a schema-driven process of global feature detection. It is important to note however, that this can only be maintained if the FFD-differences can unambiguously be interpreted as a direct reflection of this qualitative change in pattern recognition processes. This because the finding of superior memory for details of improbable objects does not constitute a sufficient basis for inferring such a change. Indeed, Friedman also found that subjects were less successful in recognizing distractor scenes in which probable objects had been completely deleted or substituted by another object rather than merely altered in some detail. While this suggests that scene information is memorized in reference to a schema-like representation of that scene (i.e. episodic memory for a given scene appears to include only its general theme and those scene elements which deviate from the already stored schema), it also indicates that, regardless of the amount of detail in which a probable object has been patten analyzed, this information is less likely to enter episodic scene memory than is the case for improbable objects. It follows then that if one wishes to maintain that this study clearly supports the notion of schema-mediated qualitative differences in the pattern recognition processes underlying probable and improbable object identification, one should be able to regard the FFD-differences as sufficient proof of this. This however, does not appear to be a self-evident matter since Henderson, Pollatsek, and Rayner (1987, 1988) reported a series of experiments suggesting a quite different explanation of these FFD-differences.

Specifically, these authors found that foveal viewing of a single object prior to making a saccade to a semantically related extrafoveally located target object, facilitated the target's identification (as measured by both naming latency and FFD) when it in turn was fixated. Clearly, since this effect was observed using arrays of isolated objects rather than coherent, expected scenes, Friedman's theory can not adequately explain its appearance. Rather, Henderson et al. propose, the effect can be interpreted as reflecting the operation of an automatic object-to-object priming process, i.e. a spreading of activation in a network of individual object representations.

It is clear that this priming mechanism is in need of further specification. Specifically, it still remains to be determined at which level of representation it operates (the distinction Kroll and Potter (1984) make between a form-specific object lexicon and an amodal conceptual store seems particularly relevant here). In addition, it is not clear yet what the precise nature of its influence is, i.e. does it primarily affect visual object processing or accessing of the object's conceptual identity? In spite of these theoretical questions however, there are sufficient grounds for assuming that the inter-object priming mechanism could very well serve as an alternative to Friedman's explanation of the FFD-differences she observed in full scenes.

First, as Henderson et al. point out, it seems quite reasonable to argue that objects appearing in the same scene tend to be semantically related. In addition, there appears to be evidence for the idea that a pattern of consecutive fixations on different objects (which, judging from the Henderson et al. research, is a necessary condition for obtaining inter-object priming) is not an exclusive characteristic of visual exploration in arrays of isolated objects. That objects also constitute the perceptually most relevant fixation locations in natural scenes, has been suggested in research by Antes, Singsaas, and Metzger (1978) and Metzger and Antes (1983). These authors demonstrated that object processing is quicker and more complete when the objects are brought into foveal vision, while the processing of setting and background information is most efficient in extrafoveal vision. In view of these considerations, it does not appear to be unreasonable to raise the question whether priming at the level of individual object representations rather than generating object identity hypotheses at the level of global scene schemas, underlies the shorter identifcation times Friedman reported for probable objects.

A study which seems to provide evidence directly relevant to this question is presented by Antes and Penland (1981). In this experiment a direct comparison was made between eye movement patterns of subjects looking at a full scene ('high context' or HC-condition) or at a 'low context' (LC) version of that scene, i.e. at an array of isolated objects constructed by simply removing all background and some of the objects present in the full scene. Two of the objects appearing in both HC- and LC-displays had a priori been rated as being highly improbable in the HC-display and were designated as improbable targets. From the remaining objects appearing in both displays (which all had been rated as probable) two were designated as probable targets. Subjects saw each display for 4 seconds in preparation of an object recognition test immediately following each display.

Guided by Friedman's theory and the additional assumption that complete and coherent natural scenes contain global contextual information which is extracted very rapidly and provides immediate access to the corresponding scene schema, Antes and Penland made two predictions which may be directly relevant to the schemas versus priming issue. First, they expected the global contextual information in the HC-displays to lead to scene schema activation. As a result, FFD's on probable targets in those displays were predicted to be shorter than those on the same objects in the LC-displays, where no contextual information was available to activate the appropriate scene schema. Secondly, they predicted that in the HC-displays only, saccades towards probable targets would be longer than those towards improbable targets, reflecting a greater useful field of view for the probable objects. Since Antes and Penland do not clearly outline the rationale underlying this second prediction, I assume that they based it on the idea that the detection of global object features specified in an activated scene schema, can occur further in extrafoveal vision and more compellingly suggests the presence of an object than is the case for data-driven recovery of detailed object features. Consequently, since objects appear to be the preferred fixation locations in natural scenes, one could indeed argue that probable objects will elicit saccades over a greater distance than improbable objects. Obviously, such an effect should not be expected in the LC-displays where scene schema activation is assumed to be absent.

An analysis of eye movement patterns showed no significant probability-related FFD-differences in either the LC- or HC-displays, but did reveal that while the FFD's for improbable targets were unaffected by context, those for probable targets were significantly shorter in the HC- than in the LC-displays. Secondly, it was found that mean saccadic amplitude was not affected by target probability in the LC-displays, but was significantly greater for probable than for improbable targets in the HC-displays.

At first sight, these findings appear to be quite consistent with the notion that global scene schema activation rather than inter-object priming underlies probability-related differences in ease of object identification in scenes. A first observation causing problems for the priming approach seems to be the absence of a priming effect in the LC-displays. Since the Henderson et al. (1987, 1988) findings seem to be reliable -priming between semantically related objects surfaced in each of their experiments as well as in previous research (e.g. Huttenlocher & Kubicek, 1983)-, this seems to suggest that the assumption that objects appearing in the same scene are generally semantically related is incorrect. However, before concluding from this that inter-object priming has no explanatory validity with regard to the context-sensitivity of object identification in scenes, there is one important consideration to be made. Specifically, it appears to be quite possible that, due to differences in the makeup of LC- and HC-displays, priming of probable targets only occurred in the latter type of displays. As a result of a much smaller total number of objects and a substantially larger proportion of improbable objects in the LC-displays, the frequency with which the fixation of a probable target was preceded by a fixation on another probable object can safely be assumed to have been considerably lower in the LC- than in the HC-displays. Obviously, this could within the confines of the priming model, explain the FFD-differences observed in this study.

As for determining the extent to which the saccadic amplitude findings necessitate the conclusion that global scene schemas drive individual object identification in scenes, the situation is somewhat more complicated.

First, it is not entirely certain whether the greater saccadic amplitude for probable targets in the HC-displays does indeed indicate that the presence of these objects can in general be detected at greater distances. Indeed, the effect does not appear to be very reliable since, in a study very similar to that by Antes and Penland (1981), Loftus and Mackworth (1978) did not find any probability-associated differences in mean saccadic amplitude prior to target fixation. According to Antes and Penland however, these findings do not challenge their theory. They point out that the absence of a saccadic amplitude difference could very well be an artefact attributable to the fact that the improbable objects in the Loftus and Mackworth (1978) stimuli tend to 'stand out' more than the probable objects. Judging from the stimulus example Loftus and Mackworth present, this could indeed be the case since here the improbable target (i.c. an octopus consisting exclusively of curvilinear line segments) is clearly more visually dissimilar to its context (a farm scene predominantly made up out of straight lines at sharp angles) than its probable counterpart (i.c. a tractor). Consequently, it seems that there is indeed some ground for arguing that in this study easily detectable low level physical discrepancies between improbable objects and their context may have compensated for their inferior detectability in extrafoveal vision, which is supposedly demonstrated in the Antes and Penland (1981) experiment.

A second question then is, whether inter-object priming could account for these detectability-differences as well as schema theory does. In this respect, I do not consider the absence of saccadic amplitude differences in the LC-displays to constitute evidence against the plausibility of such an account. Apart from the fact that it is not clear whether any priming occurred here, there also seems no reason to expect saccadic amplitude differences in this condition, even if priming were to allow for primed object detection at greater distances than is the case for unprimed objects. This because in arrays of isolated objects, every bit of extrafoveal information suggests the presence of an object (and consequently a perceptually relevant fixation location), regardless of the degree to which it is compatible with activated object representations.

As for the saccadic amplitude differences in the HC-displays, there is some evidence that inter-object priming could account for them. Using gaze duration as a measure of target identification time in arrays of isolated objects, Henderson, Pollatsek, and Rayner (1988) tested explicitly for the overadditive effect of foveal prime and extrafoveal target preview which one would expect if priming were to enhance extrafoveal target processing. Their results revealed that having a related object in the fovea did indeed enhance the amount of facilitation derived from an extrafoveal preview beyond what can be expected on the basis of a mere additivity of priming and preview effects. Admittedly, a comparison of these results with those of a similar analysis in their first series of experiments (Henderson et al., 1987), indicates that this enhancement of extrafoveal target processing is contingent upon extensive processing of the foveal prime (i.c. explicit identification and memorization). Consequently, one could argue that this finding may not reflect a process characteristic for all situations involving real-world scene perception. However, it does clearly indicate that inter-object priming could be at the basis of the saccadic amplitude differences in the Antes and Penland (1981) study, in which subjects were explicitly required to memorize scenes in anticipation of an object recognition test.

Conclusion

Based on the above discussion, it appears to be a reasonable conclusion that none of the evidence reported here, unequivocally indicates that context-consistent object identification in scenes is based on a concept-driven detection of global object features specified in a global scene schema activated prior to or during the earliest stages of scene perception. Indeed, all of the observed effects can equally well be accounted for by the assumption that data-driven access to an object representation primes related object representations, thus reducing the thresholds for establishing a match between them and the object features recovered in a further data-driven analysis of the image.

While it is clear that this priming process still needs some further specification and research, it does seem to hold the promise of a simpler and more powerful account of context effects on object identification, than what can be provided by a theory centered around the notion of global scene schemas. Indeed, it avoids the problem -posed by the Antes and Penland (1981) findings- of having to outline a theory on how the appropriate scene schema is very rapidly activated in the absence of clear scene expectations (see Biederman (1981, 1988) for some speculative notes on this topic). In addition, it accounts for context effects in both natural scenes and arrays of isolated objects, enlarging its explanatory scope relative to that of schema theory.

However, before this can be taken to provide sufficient grounds for entirely dismissing the schema approach to this domain, it is necessary to examine a second strain of research claiming that global scene-schemas play a major role in the identification of individual objects in scenes. Specifically, the following section of this chapter will be devoted to a discussion of the possibility that object identification in scenes does not only involve pattern recognition of the object itself, but is also based on the use of global scene schemas which allow for the extraction of object-diagnostic information from an object's spatial relations to the scene it appears in.

This discussion is important since it is quite clear that, contrary to what is the case for an approach centered around global scene schemas, inter-object priming can not explain effects of spatial contextual information on object perception. Consequently, any evidence for context effects of spatial scene-structure would invalidate inter-object priming and endorse scene schema ativation as the central notion in a complete account of the context-sensitivity of object identification in real-world scenes.

1.2 <u>The Role of Spatial Object-Context Relations in Object Identification in Scenes</u>

Based on an analysis of what characterizes the appearance of objects in coherent natural scenes, several authors have argued that like an object's pattern characteristics its relation to the scene it appears in provides information about the object's identity (Biederman, 1981; Biederman, Mezzanotte, & Rabinowitz, 1982; Klatsky, Teitelbaum, Mezzanotte, & Biederman, 1981).

First, they point out, the appearance of objects in scenes will typically reflect their fundamental physical nature, i.e. the fact that they are entities with a certain mass and density. Their resulting susceptibility to gravity and incapability to occupy the same position their surroundings occupy, is directly evident in a general tendency for objects to appear supported by some surface and cause occlusions in the scene they appear in. Consequently, Biederman and his colleagues argue, two object-context relations can be identified (i.e. <u>Support</u> and <u>Interposition</u>), which for any object in any natural scene provide information with regard to its general physical identity.

Secondly, these authors claim, the appearance of objects in natural scenes has several additional characteristics which allow for the definition of three more object-context relations which also hold for any object in any coherent scene, but provide more specific information pertaining to the semantic identity of objects. Specifically, the characteristics involved are i) the fact that objects are typically found in some scenes and not in others, ii) the tendency for objects to occupy privileged positions in the scenes they are likely to be found in, and iii) the existence of typical and stable size ratios between objects appearing in the same scene. One can therefore argue that Probability, Position and Size relations can be defined which allow for the formation of hypotheses about an object's conceptual identity on the basis of global and/or local interpretations of the scene it appears in (i.e. interpretations concerning the scene's global theme and/or the identity of other objects appearing in it).

Clearly, this characterization of the appearance of objects in natural scenes suggests that the research discussed in the first section may only have captured part of the context-sensitivity of object identification by exclusively focusing on the effects of purely *conceptual* aspects of context (i.e. individual object probability in a scene or semantic relatedness between consecutively attended objects) on object pattern recognition. Specifically, the question is raised here whether the *spatial* structure inherent in natural scenes does not provide a contextual definition of an additional set of relational object features (i.e. Support, Interposition, Size and Position) which are taken into account during object identification.

Before examining the evidence relevant to answering this question, it should be pointed out that the distinction between conceptual and spatial contextual aspects, should not be confounded with the physical-semantic distinction Biederman et al. proposed. The former distinction separates Probability from Support, Interposition, Size and Position on the basis of differences in the source this contextual information is drawn from (i.e. the scene's conceptual interpretation versus the scene's spatial layout). The latter distinction however, separates Support and Interposition from Probability, Size and Position on the basis of a difference in the kinds of preliminary scene processing required to use these relations as a basis for object identification. Specifically, this distinction reflects the fact that while object size and position can be encoded from a scene prior to its semantic interpretation, they only (like object probability) become distinctive object features by virtue of the object's presence in a particular scene. Consequently, this information requires a semantic interpretation of the scene in order to be used as a basis for object identification. As for support and interposition, this is not the case since they are characteristic for any object, no matter what scene it appears in. As a result, they merely need to be determined in a physical 3D-parse of the scene in order to reveal the aspect of object identity they carry information about (i.e. its fundamental physical nature). As will be explained below, Biederman et al.'s use of this distinction plays a crucial role in the theoretical conclusions they have drawn from the experiments which I will discuss presently.

In these experiments (Biederman, 1981; Biederman, Mezzanotte, & Rabinowitz, 1982), subjects were asked to determine whether a pre-named target object had been present at a specific position in a tachistoscopically presented line drawing of a scene. Scenes were exposed for 150 msec and were followed by a mask containing a dot, which indicated the position of the object the subject had to decide about whether or not it was the pre-named target. On half of the trials the target did indeed appear at the cued position while on the other half some other object was presented there. The variable of interest was the degree to which the appearance of the object at the cued position conformed to the five object-context relations defined above. In a Base condition the object violated none of its typical relations to the scene it was presented in. In various Violation conditions however, one, two or three of five possible infractions on these relations (i.e. the object floated, passed through its background, was improbable, appeared in an inappropriate position or size), were introduced.

According to Biederman et al., speed and accuracy of the subject's responses in this experiment can be regarded as a measure of the perceptibility of the object at the cued position. Consequently, they claim, the finding of violation costs in any of the Violation conditions (i.e. a decrease in response speed and accuracy relative to the Base condition) would indicate that the violated relations constitute part of the information normally used in identifying objects in scenes.

Against the background of this logic three main results have been obtained in these studies. First, all manipulated object-context relations appeared to provide contextual information used for object identification in scenes since violation costs were incurred for each one of them (with the exception of Interposition which produced no violation costs at all). Second, as more pieces of misleading contextual information were introduced, the perceptibility of the object they pertained to decreased. This was suggested by a clear increase of miss rates and correct reaction times along with a very slight but significant increase in false alarm rates, as the number of simultaneous relational violations went up from zero to three. Third, based on an inspection of the relative size of the violation costs incurred for the various types of violations, it was found that i) with the exception of Interposition, spatial contextual information (i.e. Support, Size and Position) had at least as much of an effect on object identification as conceptual contextual information does (i.e. Probability), and ii) physical relations (i.e. Interposition and Support) do not have a stronger effect on object identification than semantic relations do (i.e. Probability, Size an Position).

Much along the lines laid out by Friedman (1979) and Antes and Penland (1981), Biederman and colleagues have interpreted these results as indicating that the first 150 ms of scene viewing are sufficient to activate a global scene schema, which contains an integrated representation of both the conceptual and spatial structure of the scene. Based on this schema activation, object identity hypotheses are generated which are verified in the viewed scene by means of a search for the spatial and featural characteristics the schema specifies for each object it includes. To the extent that scene and schema information are compatible, this will result in rapid and accurate object identification, while incompatibilities between them will cause the object identification process to be slower and more error-prone.

As was already pointed out, the evidence for effects of spatial violations clearly seems to favor this theory over an inter-object priming account of context effects on object identification in scenes. In addition, Biederman and colleagues claim that the effects of semantic violations characterize most prominent models of visual object perception (e.g. Guzman, 1969; Hoffman & Richards, 1985; Marr, 1978 and 1982; Waltz, 1975) as inadequate for modeling more than the perception of un-anticipated, isolated objects. Specifically, they argue that these theories are flawed in their description of object perception as exclusively based on the data-driven and pre-conceptual recovery of the object's structural features from the image. If this were indeed a correct view, Biederman et al. point out, one could perhaps expect effects of the physical violations since they might interfere with parsing the object from the image, but one certainly should not find effects of semantic violations. Since the results showed no effects of Interposition and only a small effect of Support, while all the semantic relations did produce substantial violation costs, they conclude that bottom-up accounts of object perception should be reserved for the rare cases in which object-context relations are either inappropriate or absent.

While these conclusions may appear to be quite inevitable, it should be pointed out that there are a number of problems associated with this research, which raise serious doubts about their validity.

The main problem that should be mentioned is that the response speed and accuracy recorded in this experiment may not at all reflect the perceptibility of the object at the cued position, but rather may measure the subject's degree of uncertainty in post-perceptually deciding whether this object could indeed have been the pre-named target object. Specifically, what I want to argue is that a 100-150 ms masked exposure of a scene will frequently be insufficient to succeed in a data-driven recovery of the structural features of individual objects from the image. Consequently, if object identification is primarily based on such recovery, subjects in the Biederman experiments will often have to resort to educated guesses as to whether cued and pre-named objects were one and the same. In order to guide these guesses, subjects not only have available their a priori knowledge about the pre-named target, but they can also be assumed to have at their disposal some information about the image they just saw. Specifically, as Antes, Penland, and Metzger (1981) and Antes, Mann, and Penland (1981) demonstrated, 100-150 ms scene exposures can be sufficient to get some idea of the general theme or setting depicted in the scene. In addition, while detailed structural features of the cued object may not have been recovered during the scene's exposure, this could be the case for some of its gross spatial properties (i.e. relative size and position in the scene) which have been shown to be encoded very rapidly and prior to object identity (e.g Breitmeyer & Ganz, 1976; Kosslyn, 1987; Ungerleider & Mishkin, 1982). I should immediately point out that the relative size and position I refer to, should not be confused with the semantic object-context relations <u>Size</u> and <u>Position</u> which Biederman et al. defined. I use these terms only to refer to strictly pre-conceptually detectable object characteristics (i.e. proportion of the scene's visual angle occupied by the object, its distance to the scene's ground plane, its nearness and position relative to other objects).

Based on a comparison between these two types of contextual informa-tion and their a priori knowledge about the pre-named target object, subjects can in my opinion generate post-perceptual guesses about whether or not the cued and pre-named object were the same, which will lead to the response patterns which Biederman and colleagues interpreted as reflecting variations in perceptibility of the cued object.

For the trials on which the target is present, this post-perceptual comparison will namely produce evidence against a "yes, the cued object was the target" response whenever a violation of Probability, Size, Position or Support is introduced. For instance, deriving the theme "kitchen" from a scene will increase the subject's uncertainty about deciding that some unidentified 'blob' in that scene was a wheelbarrow. An uncertainty which will increase even further when the blob occupied only a small portion of the scene, was located at a great distance of the scene's ground plane and did not appear anywhere near to another potentially support-providing surface. Consequently, one can expect that as more of these violations are introduced, the subject's uncertainty will tend to grow and he will both take more time to finally say "yes" and be less likely to respond "yes" at all (resulting in the observed increase of miss rates and correct reaction times as the number of violations goes up).

However, for the trials on which not the target but some other object was present at the cued location, a systematic change in the subject's Base condition uncertainty should not be expected in the Violation conditions. On those trials, the violations do not pertain to the pre-named target and a post-perceptual comparison of contextual information and a priori target knowledge should therefore be largely non-informative and irrelevant to task performance. Consequently, what one would expect is an essentially identical performance level for the catch trials across the Base and Violation conditions. Note that Biederman et al.'s (1982) finding of a very slight but significant increase in false alarm rates as more violations are introduced, can hardly be viewed as a serious argument for rejecting the post-perceptual interpretation of the data in favor of an exlanation in terms of object perceptibility. Indeed, apart from having failed to replicate this finding (Klatsky, Teitelbaum, Mezzanotte, & Biederman, 1981), Biederman and colleagues are equally unable to account for it since there is no obvious reason why, within the framework of their theory, one would expect subjects to be more likely to claim that an object is a pre-named target as the object in question becomes less perceptible.

It follows from this discussion that the post-perceptual comparison explanation I proposed here, deals with the main aspects of the Bieder-man data equally well as schema theory does. In fact, when we consider some of the more detailed results it turns out that it may even be preferable.

A first argument to this effect is that Biederman et al. found that false alarm rates were consistently higher for catch trials on which the pre-named target was probable to appear in the scene, than for catch trials on which this was not the case. Obviously, this is completely in line with the post-perceptual comparison hypothesis while it poses problems to the Biederman et al. interpretation. Specifically, this finding implies that knowledge associated with the individual target named before scene exposure plays an important role in determining the subject's response. The question then becomes to what extent one can still maintain that responses in this experiment reflect influences of knowledge contained in a global scene schema activated during the very first stages of scene viewing.

Secondly, there is the total absence of an effect of Interposition violations. Following the Biederman et al. logic, this implies that a violation which thoroughly disturbs an object's featural structure has no effect whatsoever on that object's perceptibility. While one could certainly argue, as Biederman does, that this only lends additional support to the notion that in scenes relational object features play a much more important role in object identification than structural object features do, I feel rather hesitant in accepting this interpretation. The problem is that it presupposes that relational object features are generally sufficient to uniquely and correctly specify an object's identity, which I think is at the least a questionable assumption. The post-perceptual explanation however, predicts this absence of <u>Interposition</u> effects without assuming this. Specifically, it starts out from the idea that 100-150 ms scene exposures are generally insufficient to recover an object's featural structure, which logically entails that disturbances of this structure should have little effect on later decisions concerning the object's identity. I therefore see an additional reason here to be more favorable towards this explanation of the Biederman et al. results.

A final problem that should be mentioned, concerns the effects that were obtained for the multiple Violation condition which included simultaneous violations of <u>Probability</u> and <u>Size</u>. Specifically, the problem is that in this condition violation costs were higher than those obtained for the condition in which <u>Probability</u> only was violated. Within the framework of Biederman et al.'s theory this is quite an inexplicable finding since it implies that a global schema pertaining to a specific scene contains knowledge about the typical size relations that hold between that scene and all objects that typically do not appear in it. Obviously, this is a rather unlikely situation and an alternative explanation needs to be offered. Clearly, the post-perceptual comparison hypothesis is a plausible candidate here. Indeed, even if apprehension of the scene's global theme suggests that the target was improbable to be in it, subjects will still be able to determine whether the relative visual angle occupied by the blob at the cued position conforms to what can be expected if the target should be placed in that particular scene.

Conclusion

Based on the above discussion of the research by Biederman and colleagues, it appears safe to conclude that no irrefutable evidence has been presented for the theory that spatial scene-structure has a perceptual effect on object identification in real-world scenes. Consequently, the Probability effect discussed in the first section of this chapter thusfar appears to be the only reliable indication of contextual effects on object identification.

Two alternative explanations have been offered for this phenomenon : one centered around mandatory top-down influences originating in a rapidly activated global scene-specific schema; and one based on an automatic priming process operating between individual representations of semantically related objects, thus reducing thresholds for data-driven pattern recognition and identification of primed objects in the image. While it was indicated that the latter explanation seems to be preferable because of its greater simplicity and generality, a more direct test of its sufficiency as an account of the Probability effect in scenes is clearly in order. Specifically, the remainder of this thesis will focus on examining some possible constraints on inter-object priming which, if proven to be essential to this process, could serve as a basis for testing its validity as a mechanism for explaining scene-context effects on object identification.

CHAPTER 2

CONSTRAINTS ON INTER-OBJECT PRIMING

Based on the Henderson et al. (1987, 1988) studies, three possible constraints on priming can be identified which may be relevant for determining the degree to which priming could play a central role in scene-context effects.

First, it is possible that priming, to put it in terms of Shiffrin and Schneider's (1977) distinction, only affects *controlled* and not *automatic* processing of the primed object. Indeed, in all of the Henderson et al. experiments the facilitory effect of a related prime was measured contingent upon fixation (i.e. a period of primarily controlled processing) of the primed object. Even the indications of facilitated extrafoveal processing found in the gaze durations for primed objects (Henderson et al., 1988), can be regarded as an effect on controlled object processing, since a period of selective extrafoveal attention to the primed object is very likely to have preceded that object's fixation (Morrison, 1984). This present limitation of priming observations to cases of controlled object processing is quite interesting since advocates of the schema-approach to scene-context effects (e.g. Antes & Penland, 1981) have argued that, due to their contextual facilitation, probable objects in a scene may remain entirely unattended and still can be identified through automatic feature detection. One of the objectives of the present research therefore was to examine whether priming effects are indeed constrained to controlled object processing. This was done in order to determine whether an investigation of the perceptibility of unattended objects in full scenes can be instrumental for assessing the validity of a priming account of object probability effects on object recognition.

A second possible constraint that will be examined is the apparent necessity for the prime itself to be subjected to controlled processing in order to have an effect on the processing of related objects. Using arrays of four isolated objects, Henderson et al. (1987) found the decrease of the first fixation duration on a target object to be strictly conditional upon the immediately preceding fixation of a target-related object. The mere presence of other target-related objects in the array yielded no such effect as evidenced by the absence of a decrease in first fixation duration when the target was either the first object to be fixated in the display or was fixated following the fixation of an unrelated object. Additional support for this absence of a 'display-consistency effect' was reported by Henderson et al. (1988) who found that an object's semantic relatedness to the array it appeared in did not affect the amount of benefit derived from its extrafoveal preview.

As was the case for the first constraint, this apparent dependency of priming effects on controlled prime processing suggests that priming may be insufficient to account for context effects in full scenes. Specifically, if this second constraint does indeed hold, a strict priming view would predict that the facilitation of an object's processing in a scene should be a function of its semantic relation to the previously fixated object while its relation to the rest of the scene should be of little importance. Concretely, this would mean that no facilitation should be expected for probable objects which are fixated as the first object in a scene or are fixated following the fixation of an improbable object. The first of these predictions runs counter to the Klatsky, Teitelbaum, Mezzanotte, and Biederman (1981) claim that a 100 ms scene exposure is sufficient to produce facilitation for a probable object foveated during that exposure. The second prediction is contested by the Antes and Penland (1981) suggestion that probable objects can be identified extrafoveally even when an improbable object is being fixated. Clearly, if priming could be demonstrated to be strictly conditional upon controlled prime processing just prior to target processing, an unambiguous confirmation of these two claims would indicate the insufficiency of inter-object priming as an account of scene-context effects. Taking into consideration that other authors have claimed that priming can be initiated on the basis of automatic prime processing (e.g. McCauley, Parmelee, Sperber, & Carr, 1980), it therefore was decided to examine the reliability of the Henderson et al. (1987, 1988) failure to observe display-consistency effects.

The third possible constraint that will be investigated is more than likely the most important one. Specifically, probably the main objection which one could formulate against the idea that priming might underly the facilitation of probable objects in scenes, is that the results presented as proof for priming were obtained with groups of objects selected on the basis of their *semantic relatedness*. While Henderson et al. (1987) assume (and quite reasonably I think) that a probable object in a given scene is more likely to be semantically related to the other objects in it than is the case for an object which is improbable in that scene, one can undoubtedly come up with an impressive list of non-related objects which are likely to appear in the same scene (e.g. a toilet and an electric razor in a bathroom, a fireplace and a television in a living-room, etc.). In fact, as pointed out in the discussion of the Antes and Penland (1981) experiment in the first chapter, it is not impossible that the absence of facilitation for probable objects in the Low Context condition of that study should be interpreted as showing that the priming effect is indeed strictly limited to objects that are clearly semantically related. Clearly, this suggests that an orthogonal manipulation of an object's probability in a scene and its semantic relatedness to the other objects in the scene could provide the necessary data for determing whether priming plays any role in object probability effects in scenes. In order to establish whether this would be a useful strategy, a third objective of the present research was to systematically examine the existence of priming effects between objects selected on the basis of their common likelihood to appear in a given scene, i.e. on the basis of their *episodic relatedness*.

CHAPTER 3

EXPERIMENTAL RESEARCH

3.1 Experiment 1

3.1.1 Research objectives and approach

The main objectives of this experiment were to determine whether inter-object priming could affect automatic processing of a primed object and if so, whether controlled processing of the prime would be a necessary condition for this effect to appear. In other words, this experiment is an attempt to establish whether the claim (Antes & Penland, 1981; Biederman et al., 1982) that identification of a probable object in a full real-world scene is facilitated regardless of whether it or other probable objects are attended to, could in principle be explained as the result of an inter-object priming mechanism.

To the best of my knowledge, only one study has been reported which provides evidence that appears to be directly relevant to these issues. Specifically, in order to determine the relative ease with which global scene and individual object information are processed during the first glance at a scene, Antes, Penland and Metzger (1981) measured accuracy of a target object's recognition in a forced choice task following its 100 msec presentation in either a full scene (High Context) or in an array of isolated objects (Low Context). By orthogonally manipulating the probability of appearance of target and distractors in the full scene, Antes et al. were able to determine that in the High Context condition subjects primarily responded on the basis of the scene's global theme (i.e. they showed a strong tendency to choose the objects with the highest likelihood of appearance, regardless of whether they actually had been present in the scene). In the Low Context condition, responses appeared to be primarily based on what objects the subjects had actually identified perceptually. That is, responses (1) were more accurate than in the High Context condition, (2) showed a clear superiority for objects closer to the central fixation point and (3) were unaffected by the manipulation of likelihood. It appears that two conclusions can be drawn from these results.

First, it is suggested that the global theme of a scene can be apprehended more quickly than the identity of the individual objects in it. While this finding certainly is compatible with the claims advanced in schema-theories of scene perception, it should be pointed out that it provides insufficient grounds for assuming that individual object perception in scenes is inevitably mediated by a quickly derived global scene interpretation. Following the Reicher (1969) and Wheeler (1970) rationale this could only have been inferred from the Antes et al. (1981) data if recognition accuracy of probable targets among equally probable distractors had been significantly better than recognition accuracy of improbable targets among equally improbable distractors. While the Antes et al. data appear to indicate that this was indeed the case, it should be noted that the recognition accuracy for improbable targets among improbable distractors was significantly below chance-level performance which shows major problems in the selection of distractors for this task. Clearly, this makes it impossible to interpret these data as reliable evidence for a genuine contextual facilitation of the perception of probable objects in scenes.

Second, and more relevant to the present discussion, the data obtained in the Antes et al. study suggest that here we may have the prototype of a paradigm which allows for an assessment of the effects of priming on automatic object processing. Specifically, what is of interest here is that accuracy of object recognition in the forced choice task varied as a function of factors affecting the perceptibility of the target object -i.e. degree of lateral masking and visual acuity-. This is indicated by the fact that it was higher for isolated targets than for targets presented in scenes, and also higher for targets presented closer to the central fixation point. The reason why this is interesting is that these effects surfaced following the 100 ms presentation of an uncued target at an unspecified and generally extrafoveal position. Under these conditions one can reasonably assume that subjects generally did not succeed in selectively attending to the target during its exposure and therefore could not engage in its controlled processing. This then leads to the conclusion that the object recognition accuracy recorded here is sensitive to genuine perceptual differences in the ease with which object information is acquired through automatic processing, and therefore is well-suited to measure priming effects on this kind of processing.

In fact, it could appear as if the Antes et al. (1981) data for the Low Context condition are already sufficient to conclude that inter- object priming does not affect automatic object processing since no differences in recognition accuracy for probable and improbable objects were found in that condition. However, there are two aspects of this study which make it impossible to draw this conclusion.

First, all the objects used to construct the Low Context stimuli were selected on the basis of episodic rather than semantic relatedness. Obviously, this leads to the problem that there is no way of determining whether the absence of a difference between probable and improbable objects should be interpreted as showing that priming only works between semantically related objects or simply does not apply to automatic object processing. Second, the interpretational problems are even further enhanced when one considers the fact that there was no systematic manipulation of the information appearing in foveal vision during the target's extrafoveal presentation. Consequently, it is possible that the third possible constraint on inter object priming (i.e. that it requires controlled processing of the prime) was violated as well, which makes it even more difficult to use the Low Context data as a basis for determing whether or not priming has any effect on automatic object processing.

Against the background of this analysis of the Antes et al. (1981) experiment, it was decided to use a modified version of their paradigm in order to study the role of inter-object priming in automatic object processing. Specifically, subjects were confronted with a 150 ms, masked presentation of an array of isolated objects and were then asked to indicate which one of a set of four objects had been present in the display. The presented arrays always contained 5 or 6 objects which could be grouped into two different "episodic categories", i.e. object groups defined by the common likelihood of their members to appear in the same scene. The two categories instantiated in each display were always selected so that the overlap between them could be considered to be minimal, i.e. objects selected for their high likelihood to appear in the one scene were quite unlikely to also be encountered in the other scene.

Within this basic stimulus structure, two crucial manipulations were introduced. The first manipulation concerned the nature of the *foveal information* present while the to-be-recognized object (henceforth called the target) was presented extrafoveally. In the first type of display ("Foveal Related" conditions) the foveal object was from the same category the target belonged to; in the second type ("Foveal Unrelated" conditions) the foveal object belonged to the other category instantiated in the array; and in the third type ("Foveal Absent" conditions) no foveal object was present at all. The second manipulation pertained to the nature of the extrafoveal information present during display exposure. Specifically, by varying the number of target-related extrafoveal objects in the arrays, several levels of "Extrafoveal Relatedness" were created. The rationale behind these manipulations was as follows.

First, it was assumed that a comparison of the accuracy with which the briefly and extrafoveally presented targets were recognized in the Foveal Related and Unrelated conditions, should provide information about whether *controlled prime processing* can facilitate automatic processing of a related object. Specifically, if this would indeed be the case the target should be recognized more accurately in the Foveal Related conditions. (Note 1)

Second, since the experiment also aimed at examining possible effects of *automatic prime processing*, it was decided to compare target recognition performance across levels of Extrafoveal Relatedness in Foveal Absent arrays. Finding an increase in target recognition accuracy as Extrafoveal Relatedness in these arrays increases, would obviously constitute strong evidence for the sufficiency of automatic prime processing to initiate priming effects. Unfortunately, however, the absence of such an increase would not allow for an equally clear interpretation since at least three explanations could be suggested for it. First, automatic prime processing may elicit no priming effects. Second, automatic prime processing may only have an effect conditional upon simultaneous controlled processing of another prime. In other words, the presence of unattended, extrafoveal primes may only have an effect if their perceptibility is enhanced by the identification of a foveal prime. Note that this kind of secondary priming by unattended primes has already been suggested in research on sentence processing (Paap & Newsome, 1981). Third, automatic prime processing could by itself be sufficient to elicit priming, but the total absence of foveal load in the Foveal Absent arrays could provide subjects with such a high-quality extrafoveal target preview that priming may not provide any additional benefit (a phenomenon already observed in the naming latency experiments reported by Henderson et al., 1987).

In order to unravel these possible interpretational problems, a third analysis was planned which involved a comparison of Foveal Related-Unrelated differences in target recognition performance at different levels of Extrafoveal Relatedness. Specifically, if this difference could be demonstrated to increase when Extrafoveal Relatedness in the

Foveal Related condition is increased while that in the Foveal Unrelated condition is decreased, a facilitory effect of the presence of unattended primes would be indicated. In addition, if this increased difference would prove to be the combined result of an increase in target recognition in the Foveal Related condition <u>and</u> a decrease in the Foveal Unrelated condition, automatic prime processing would appear to be self-sufficient to elicit priming. Alternatively, if the increased difference would be attributable to a target recognition increase in the Foveal Related condition <u>only</u>, unattended prime effects would prove to be conditional upon simultaneous foveal prime processing.

To conclude this general description of the experiment, a few final comments should be made with regard to the employed measure of target perceptibility.

In order to minimize the possibility for subjects to selectively attend to the target during its presentation, targets were uncued and were presented for only 150 ms at an a priori unspecified extrafoveal location in an array of objects. In addition (and contrary to the Antes et al. (1981) experimental situation), each display was followed by a visual noise mask to prevent subjects from using CRT after-images or iconic memory to turn their attention to specific extrafoveal objects in the display. This was an important control since there are indications (Loftus & Mackworth, 1978; Antes et al., 1981) that attention tends to shift very rapidly towards objects that are episodically unrelated to the scene or array of objects they appear in. In the present experiment this would systematically favor perceptual processing of the extrafoveal objects belonging to the least represented episodic category in the display, and thus would obscure any possible priming effects produced by the manipulations of target-display consistency and foveal information. While the 150 ms exposure duration was assumed to be sufficiently short in order to avoid attentional shifts *during* stimulus presentation, the presence of both a CRT after-image and an undisturbed iconic representation of the stimulus could allow for such shifts *following* stimulus presentation, which is why the mask was introduced.

Finally, following the mask, subjects were presented with a set of four object names from which they had to select the target. Contrary, to what was the case in the Antes et al. (1981) experiment, the distractors in this set always belonged to the same episodic category the target belonged to. This was done in order to ensure that responses would indeed reflect the perceptibility of individual targets rather than a general judgment about which episodic category had been represented in the display.

3.1.2 Method

Subjects

16 members of the University of Massachusetts subject pool participated in the experiment. All of the subjects had normal vision and did not require corrective lenses for reading.

Stimuli

To construct the necessary object arrays, 123 line drawings of different objects were used. A large number of these objects was drawn from the standardized set provided by Snodgrass and Vanderwart (1980). From this pool, 32 episodic categories of five objects each were assembled and divided into 16 pairs of non-overlapping categories. In each of the resulting 16 groups of ten objects, two objects (one from each category) were designated to be *extrafoveal targets*, two others (one from each category) were selected to serve as *foveal primes*, and the remaining six were assigned the role of *extrafoveal primes*. Where norms were available, targets, foveal and extrafoveal primes from the two categories in a given pair were selected to be of comparable visual complexity. A complete list of the 16 category-pairs is provided in Appendix A.

From each of these 16 pairs, eight different displays were constructed as schematically illustrated in Table 1.

Table 1. Schematic representation of structure and content of experimental display types. A1 through A5 and B1 through B5 denote different objects, with episodic membership of the object indicated by the letter and the role of the object (target, foveal prime or extrafoveal prime) indicated by the number.

DISPLAY 1a	DISPLAY 1b	DISPLAY 1a'	DISPLAY 1b'
A3 B3 A2 B1 B4 A1	A3 B3 B2 B1 A4 A1	A3 B3 B1 B4 A1	A3 B3 B1 A4 A1
DISPLAY 2a	**DISPLAY 2b**	**DISPLAY 2a'**	**DISPLAY 2b'**
A3 A5 A2 B1 A4 A1	B5 B3 B2 B1 B4 A1	A3 A5 B1 A4 A1	B5 B3 B1 B4 A1

In Table 1, alphanumerical combinations "A1" through "A5" and "B1" through "B5" represent ten different objects belonging to two episodic categories "A" and "B". The numbers in these symbols indicate the role that was a priori assigned to the object in question : "1" for targets, "2" for foveal primes and "3" through "5" for extrafoveal primes.

As seen in Table 1, the basic display structure consisted of two targets and three extrafoveal primes placed on the corners of an imaginary pentagon. In this manner inter-object and object-to-center of display distance was kept constant in order to (given the subject's fixation on the display center) equate all objects for lateral masking and visual acuity effects. Within a given pair of episodic categories, targets always appeared at the same location in order to maximize comparibility of the target's accuracy of recognition across display types. Across the 16 category-pairs, however, targets were rotated through all peripheral positions in order to ensure that the subjects would not be able to generate expectations about target positions.

Within the framework of this basic display structure the nature of the foveal information present at exposure time as well as the number of extrafoveal primes were manipulated in order to test the hypotheses outlined in section 3.1.1.

Displays 1a, 1b, 2a and 2b were constructed to examine Foveal Related-Unrelated differences in target perceptibility at various levels of Extrafoveal Relatedness. Specifically, in displays 1a and 1b, the two targets (A1 and B1) were presented in a Foveal Related condition (1a for A1 and 1b for B1) and a Foveal Unrelated condition (1b for A1 and 1a for B1) with a similar, low Extrafoveal Relatedness in both cases (i.e. one extrafoveal prime in the Foveal Related condition and two in the Foveal Unrelated condition). In displays 2a and 2b, the number of extrafoveal primes was increased from one to three in the Foveal Related condition (2a for A1 and 2b for B1), while it was decreased from two to zero in the Foveal Unrelated condition (2a for B1 and 2b for A1).

Displays 1a', 1b', 2a' and 2b' were constructed to examine whether unattended prime processing could by itself affect target perceptibility. By simply removing the central objects in displays 1a, 1b, 2a and 2b, these displays presented the targets at four levels of Extrafoveal Relatedness. In increasing order of relatedness: displays 2b', 1a', 1b' and 2a' for the A-target, and displays 2a', 1b', 1a' and 2b' for the B-target.

In this manner, the eight display types represented in Table 1 produced eight context conditions in which to be recognized targets were presented. Table 2 summarizes how the display types map onto the context conditions for the A and B-targets.

In order to measure target perceptibility in these eight context conditions, a choice set of four object names was assembled for each of the 32 targets used in the experiment. In addition to the name of the target itself, this set contained the names of three other objects that did not appear in the display the target was presented in, but did belong to the same episodic category the target was a member of. A complete list of these choice sets is provided in Appendix B.

Table 2. Context conditions for the extrafoveally presented target, produced by the variation of foveal information (target-related, unrelated or absent) and the number of extrafoveal target-related objects (ranging from 0 to 3). Display types (see Table 1) mapping onto each context condition are indicated for A and B-targets separately.

FOVEAL RELATED / EXTRAFOVEAL 1	FOVEAL UNRELATED / EXTRAFOVEAL 2	FOVEAL RELATED / EXTRAFOVEAL 3	FOVEAL UNRELATED / EXTRAFOVEAL 0
A-targets : - display 1a B-targets : - display 1b	A-targets : - display 1b B-targets : - display 1a	A-targets : - display 2a B-targets : - display 2b	A-targets : - display 2b B-targets : - display 2a
FOVEAL ABSENT / EXTRAFOVEAL 0	**FOVEAL ABSENT / EXTRAFOVEAL 1**	**FOVEAL ABSENT / EXTRAFOVEAL 2**	**FOVEAL ABSENT / EXTRAFOVEAL 3**
A-targets : - display 2b' B-targets : - display 2a'	A-targets : - display 1a' B-targets : - display 1b'	A-targets : - display 1b' B-targets : -display 1a'	A-targets : - display 2a' B-targets : - display 2b'

Finally, in addition to the experimental stimuli, 24 more arrays of six objects each were constructed to serve as practice and filler stimuli. While the spatial structure of these displays was identical to that of the experimental stimuli, the six objects in them always belonged to six different episodic categories and the objects designated as targets were always located in the center of the display. There were two reasons for introducing these arrays as practice and filler stimuli. First, their categorical heterogeneity could help to discourage subjects from responding on the basis of general categorical display-consistency impressions instead of on whether or not they actually saw any of the choice alternatives. Second, the central location of the targets in them could keep the subjects from adopting the strategy of covertly shifting their attention to the peripheral regions of the displays while neglecting the central area (which would obviously defeat the purpose of the experiment). Complete lists of these practice and filler stimuli and their corresponding sets of choice alternatives are provided in Appendices C and D.

Apparatus

The object pictures and the mask were entered into a Hewlett-Packard 2100 computer by means of a Summagraphics Bit-Pad, and were displayed on a Hewlett-Packard 1300A CRT with a P-31 phosphor. Over the entire set of objects, pictures subtended from 1° to 3° both horizontally and vertically, while the mask subtended 4° by 4°. Inter-object distance (measured from center to center) and object-to-display center distance were approximately 6°. The sets of four object names presented as alternatives in the forced choice target recognition task, were also displayed on the CRT (one below the other), with the target name's position rotated through the list across trials.

Procedure

Upon arrival, subjects were seated 46 cm from the CRT with their head held in position by a chin and forehead rest in order to keep viewing distance constant and eliminate head movements. Subjects were told that they would see a series of brief presentations of object groups, which all would be followed immediately by the presentation of a set of four object names. One of these names, they were told, would always correspond to an object that had actually been present in the display they just saw, while the other three would not. Their task then was to pick out the name of the object that had indeed been in the display. Each subject received a total of 56 trials : 4 practice trials and 32 experimental trials with the 20 filler trials inserted at fixed positions between them. Each trial consisted of the following events: First, a cross was presented in the center of the display and the subjects were instructed to fixate it. Subsequently, an array of objects was displayed for 150 ms, immediately followed by a 250 ms presentation of the mask at each of the locations where an object had just appeared. Following the offset of the mask, a set of object names were displayed as choice alternatives and the subject's choice was recorded by the experimenter. No feedback was given before the end of the experiment which lasted 15 to 20 minutes.

Design

As pointed out above, the measure of interest in this experiment was the accuracy of target recognition in each of the eight context conditions presented in Table 2. Two restrictions had to be taken into account in obtaining this measure. The first one was that subjects could not be presented with the same array of objects more than once, since previous research on the effects of extrafoveal primes (Paap & Newsome, 1981) showed these effects to be a function of the subject's familiarity with the stimuli. The second one was that subjects could not be presented with the same set of choice alternatives more than once, in order to avoid effects of response strategies (e.g. a subject could attempt to show consistency by always choosing the same alternative regardless of whether or not he actually saw it). In view of these restrictions, it was impossible to present each subject with the 256 trials (16 pairs of episodic object-categories X 2 sets of choice alternatives for each of these category-pairs X 8 different displays for each category-pair), which were required to probe all targets in all context conditions. Obviously, this ruled out the possibility of measuring accuracy of target recognition in terms of the number of subjects that correctly recognized a particular target in a particular context condition. It was therefore decided to measure accuracy of target recognition for a particular context condition in terms of the number of targets in that condition that had been correctly recognized across subjects.

In order to do so, the 16 types of target recognition trials (which were produced by the combination of two sets of choice alternatives with eight different displays for each category-pair) were grouped into eight pairs of maximally discrepant target recognition trials. Specifically, the two trials in each pair always differed in terms of the target that was probed for (the A-target in one trial and the B-target in the other), the foveal content of the display presented during the trial (a foveal object was present on one trial and absent in the other), and the identity of the non-target objects in the display (only one of these objects was the same in the two trials). By assigning the eight trial-pairs for a given category-pair to 8 different subjects, and repeating this procedure for all 16 category-pairs (with the restriction that across category-pairs each subject should receive all types of target recognition tasks equally often), each subject was assigned a series of 32 trials (individually randomized for each subject), while across subjects all 32 targets were probed once in each of the 8 context conditions. Since 16 subjects participated in the experiment, the whole procedure could be replicated and 64 data-points (i.e. two for each target) were available to compute the proportion correct in each of the eight context conditions.

3.1.3 Results

A first result that needs to be mentioned is that on the 20 filler trials (where the target was located in the center of the display) all subjects performed at a very high level of accuracy : proportion correct ranged from .75 to 1.00 with an average of .85 (chance-level performance being .25). Clearly, this suggests that subjects were unlikely to have adopted a systematic strategy of shifting their attention to the peripheral regions of the displays, which would have run counter to the objectives of the experiment.

The results pertaining to the accuracy of target recognition in the eight context conditions, are presented in Table 3.

Table 3. Proportion of targets correct per Context condition.

FOVEAL RELATED / EXTRAFOVEAL 1	FOVEAL UNRELATED / EXTRAFOVEAL 2	FOVEAL RELATED / EXTRAFOVEAL 3	FOVEAL UNRELATED / EXTRAFOVEAL 0
(I) .250	(II) .343	(III) .265	(IV) .406
FOVEAL ABSENT / EXTRAFOVEAL 0	FOVEAL ABSENT / EXTRAFOVEAL 1	FOVEAL ABSENT / EXTRAFOVEAL 2	FOVEAL ABSENT / EXTRAFOVEAL 3
(V) .421	(VI) .437	(VII) .390	(VIII) .406

The first thing to note here is that the comparison between the Foveal Related (I and III) and Unrelated (II and IV) conditions does not show superior target recognition performance in the former, as would be expected if controlled prime processing facilitated automatic perceptual processing of related objects. In fact, the difference was clearly in the opposite direction since the overall proportion correct in the Foveal Related conditions (i.e. .257) is significantly smaller [$\underline{Z} = 2.01$, $\underline{p} < .05$] than that in the Foveal Unrelated conditions (i.e. .374).

Contrary to this clear indication of an, albeit unexpected, Foveal Relatedness effect on target recognition performance, no such indication appears to be present for an Extrafoveal Relatedness effect. In the data for the Foveal Absent conditions (proportions V through VIII), no significant differences could be found in the six possible pairwise comparisons between these conditions. Obviously, this finding does not support the hypothesis that unattended prime processing can by itself facilitate automatic processing of a related object. As mentioned before, however, a null-effect in these comparisons does not necessarily mean that Extrafoveal Relatedness can play no facilitory role at all conditional upon simultaneous Foveal Relatedness and/or the absence of a high-quality extrafoveal target preview. Since, however, an insignificant superiority of Foveal Unrelated over Foveal Related (i.e. proportion II minus proportion I, $\underline{Z}$ = 1.157, $\underline{p}$ > .05) further increases to reach significance (i.e. proportion IV minus proportion III, $\underline{Z}$ = 1.718, $\underline{p}$ < .05) when Extrafoveal Relatedness is decreased in the Foveal Unrelated condition while it is increased in the Foveal Related condition, the data quite clearly argue against any facilitory effect of unattended primes on automatic processing of related objects. To the contrary, the data for the Foveal Unrelated conditions show a tendency for performance to decrease as Extrafoveal Relatedness increases. While this tendency does not reach significance (proportion IV minus proportion II, $\underline{Z}$ = .739, $\underline{p}$ > .05) it is interesting to note that it is consistent with a similar tendency in the Foveal Absent conditions. Only in the Foveal Related conditions this pattern did not surface, which could very well be due to a floor-effect since performance in these conditions dropped entirely to chance-level. Note that this minimal level of performance was only found in the Foveal Related conditions, while recognition accuracy in all other conditions was significantly higher than chance-level (a proportion correct of .338 being sufficient to reach a .05 significance level).

3.1.4 Discussion

The first conclusion that can be drawn from the data presented here is that neither controlled nor automatic processing of individual objects appear to facilitate pre-attentive processing of other objects likely to appear in the same real-world scene. This follows from the failure to find either 1) superior target recognition in the Foveal Related versus Unrelated conditions, or 2) an increase in performance as the degree of Extrafoveal Relatedness increases. Naturally, these data do not necessarily imply that the inter-object priming mechanism observed in previous studies (e.g. Carr, McCauley, Sperber, & Parmelee, 1982; Henderson et al. 1987; Reinitz, Wright, & Loftus, 1989) is restricted to controlled processing of the target object. In order to draw this conclusion, one would have to be certain that the absence of facilitory priming effects in the present experiment was not due to the operationalisation of prime-target relatedness in terms of episodic rather than semantic relatedness. However, the data do imply that if ease of pre-attentive object processing in full scenes could be demonstrated to be enhanced by the object's likelihood of appearance in the scene, inter-object priming would provide an insufficient basis to account for this effect of scene context.

The most intriguing aspect of the data however, is not the failure to find facilitation of targets related to foveal or extrafoveal primes, but the apparent superiority of recognition for targets which were not related to these primes.

A first explanation one could propose for this phenomenon is based on the proposition that the featural dissimilarity between the target and the other objects is greater when they do not belong to the same episodic object category. Analagous to what a number of authors have suggested for semantically related objects (e.g. Carr et al., 1982; Huttenlocher & Kubicek, 1983; Sperber, McCauley, Ragain, & Weil, 1979), Biederman (1981) obtained some results indicating that episodically related objects are visually more alike than unrelated objects. Specifically, Biederman asked subjects to determine whether a pre-named target object was present in a briefly exposed (100 ms) array of extrafoveally located objects. False alarm rates in this task were lower and unaffected by the number of objects in the display only when the target did not belong to the episodic category all the other objects in the display belonged to. Miss rates however, were similar for both display-consistent and inconsistent targets and increased as the number of objects in the display increased. While Biederman initially interpreted these data as evidence for a categorical pop-out effect -similar to what Egeth, Jonides, and Wall (1972) found for letters and digits- he later (Biederman, 1982) stated that they should rather be seen as indicating that objects from the same episodic category are visually more alike and therefore more confusable than objects from different categories. In view of these results, one could argue that in the present experiment, a greater featural dissimilarity between unrelated targets and the other objects may have enhanced their visual distinctiveness. As a result, their perceptibility and subsequent accuracy of recognition could have increased directly by making their features stick out in the visual field, or indirectly by eliciting pre-saccadic shifts of attention towards these visually distinct objects.

However, two aspects of the data suggest that this explanation is not quite complete. First, while there was a slight tendency for a decrease in performance as Extrafoveal Relatedness (i.e the number of extrafoveal primes) increased, it never reached the level of significance that would have corroborated the visual dissimilarity hypothesis. Second, adding a target-related object in the center of a display did not produce a graceful degradation of target recognition performance for that display, as one would expect if the foveal prime would constitute just another source of visual target-distractor confusability. Rather, it caused performance to completely drop to chance-level, which could not be ascribed to an increase in foveal load (Ikeda & Takeuchi, 1975) since the phenomenon did not occur when a target-unrelated foveal object was added.

In view of these considerations, it appears to be more adequate to explain the unrelated target superiority by assuming that when an object's representation is sufficiently strongly activated, it will prime the representations of other objects belonging to the same episodic category. If one further assumes, that such strong activation of a single object representation and the resulting priming of related representations can only be achieved on the basis of controlled object processing, the observed results can be explained. Specifically, the chance-level performance for foveal-related targets can be attributed to insufficiently large differences in the level of representational activation between foveal-related objects that were actually in the display and primed objects that were not. Indeed, this would result in a set of possible response candidates for the subsequent forced choice task, which would include the entire set of response alternatives and therefore would necessitate the subject to choose at random. For the foveal-unrelated targets, this kind of priming and response competition between target-related objects should not be expected if the assumption is correct that controlled object processing is required for such effects to appear. The absence of a clear effect of Extrafoveal Relatedness supports this assumption and suggests that the slight tendency towards such an effect in the Foveal Unrelated and Foveal Absent conditions, may in fact be attributable to the identification of (and subsequent priming by) extrafoveal target-related objects on a small proportion of the trials.

In conclusion, the present data are taken to, albeit indirectly, suggest that inter-object priming is not limited to semantically related objects but also occurs between episodically related objects. However, the results also indicate that if such priming exists, it does not facilitate pre-attentive processing of primed objects. This clearly suggests an approach for unambiguously testing the validity of inter-object priming as a complete account of effects of object probability in full scene context.

3.2 Experiment 2

3.2.1 Research Objectives and Approach

The first objective of this experiment was to further explore the superior recognition of unrelated targets in the Foveal Present conditions in Experiment 1. Specifically, the question was examined whether this finding reflected a greater perceptibility of these targets or rather should, as proposed in the discussion of Experiment 1, be interpreted as indirect evidence for the operation of a priming process between episodically related objects.

As mentioned before, the absence of systematic effects of Extrafoveal Relatedness in the first experiment argued against the hypothesis that a 'pop-out' effect based on categorical (Egeth et al., 1972) or featural (Biederman, 1982) dissimilarity could have directly enhanced the perceptibility of the unrelated targets. However, one could still stress the point that the unrelated target recognition was only significantly superior in the displays containing no other target-related objects (i.e proportion IV in Table 3). Consequently, under the assumption that objects from different episodic categories do indeed tend to be more visually dissimilar, one could argue that this indicates the possibility of a pre-saccadic attentional shift towards those elements in the periphery of the display which are visually most dissimilar to the information processed foveally. On this view, one would indeed expect the unrelated targets to be recognized most accurately in displays in which such an attentional shift would systematically be directed towards them, i.e. in the displays in which all other peripheral objects are related to the foveal object.

In order to determine the validity of this alternative explanation of the unrelated target superiority, it was decided to examine one of its possible implications in the present experiment. Specifically, if featural dissimilarity between the object in foveal vision and a peripherally located object were indeed to draw the subjects' attention, one could expect this attentional shift to be followed by a saccade towards this dissimilar object. In fact, research by Loftus and Mackworth (1978), Antes and Penland (1981) and Antes et al. (1981) does indeed suggest that both in full scenes and in arrays of isolated objects, peripherally located objects tend to be fixated earlier in the course of display exploration when they are episodically unrelated (and hence featurally dissimilar) to the objects foveated during the first fixations on the center of the display. Consequently, it appeared to be worthwile to examine the scanning patterns of subjects confronted with the Foveal Present stimuli of Experiment 1, in order to determine whether the unrelated target superiority observed for these stimuli could be attributed to covert attentional shifts towards these targets.

Clearly, if this analysis of scanning patterns should reveal that unrelated objects in the periphery of the displays tend to be fixated earlier, the related-unrelated differences in Experiment 1 could no longer be interpreted as a result of response competition caused by a priming process between episodically related objects. Therefore, the second objective of Experiment 2 was to more directly examine the existence of such priming by using a measure generally believed to directly reflect ease of object identification (i.e the duration of the first fixation on an object). Obviously, a decrease in first fixation duration for a given target object following the fixation of an episodically related object would lend support to the notion of a priming process between these objects.

In addition, first fixation duration for targets can also be analyzed as a function of the simultaneous presence of other target-related objects which are not attended to prior to the targets' fixation. In this manner, it can be established whether the failure for automatic prime processing to facilitate target recognition in Experiment 1, should be attributed to the pre-attentive nature of the prime processing or of the target processing in that task. The primary reason for examining this question was that its answer can provide yet another step towards identifying the kind of research that is most likely to yield unambiguous conclusions with regard to the validity of an inter- object priming explanation of probability effects on object perception in scenes. Specifically, if the presence of unattended primes would have no influence on first fixation durations for the targets, then a priming account of object probability effects in scenes would be clearly falsified by any effects that are not exclusively attributable to the episodic relation between the target and the object fixated just prior to it. In other words, this would lead to the very concrete prediction that no facilitation should be found for probable targets fixated as the first object in a scene, or for probable targets fixated after attending to an improbable object.

It may seem that the last question examined in the present experiment has already been answered in the Henderson et al. (1987, 1988) studies where global target-display consistency failed to have an effect on the duration of first target fixations. There is one important reason, however, why these data do not conclusively rule out the possibility of an effect of automatic prime processing. Specifically, as Henderson et al. (1988) already mentioned, an effect of automatic prime processing presupposes that visual attention can be unfocused to such an extent that object information can be extracted across the entire display. As Loftus (1983) argues, this widely distributed attention may well be a typical characteristic of the first fixation(s) on a scene, allowing for the detection of some object features and their spatial locations. While this proposition finds some support in the cited evidence for very rapidly occuring systematic saccades towards peripherally located inconsistent objects, Henderson et al. (1988) argue that their data favor an alternative model of visual attention. Specifically, while they did find indications for an extrafoveal preview benefit for targets fixated following the fixation of a related object, no such benefit was observed when the target was related to all the objects in the display except the one fixated just prior to it. Consequently, they claim, a sequential model of visual attention seems more appropriate, according to which attention is only directed towards the position currently being fixated and the one about to be fixated.

The question which can be raised, however, is whether Henderson et al. did not induce their subjects to distribute their attention in such a sequential fashion, by imposing a standard fixation sequence of objects appearing at fixed positions in displays terminated by the subjects themselves. It appears to be a reasonable assumption that under these conditions, there is no need for a wide distribution of attention and consequently little reason to expect target-display consistency effects. A fairer test for the possible existence of effects of automatic prime processing therefore was required, in which task conditions explicitly favored a wide distribution of attention.

In order to examine these questions a modified version of the paradigm employed in Experiment 1 was used. Using similar arrays of isolated objects, subjects now were allowed to move their eyes away from the center of the display and fixate one of the peripheral objects. Displays were terminated and the mask and forced choice task were presented once a peripheral object had received a first fixation, or if no such fixation had occurred after 1 sec of exposure. These exposure time constraints were imposed in order to induce the subjects to actually move their eyes away from the display center to the peripheral objects. In addition, they forced subjects to gather spatially disparate information and make a fixation decision in a limited amount of time, thus favoring an initially wide distribution of attention. During the entire exposure duration of the displays, eye movement patterns were recorded in order to control display termination and collect the necessary data for testing the hypotheses outlined above.

Specifically, scanning patterns across displays with a central object should reveal whether peripheral objects which are not episodically related to that object are selec-

ted for fixation more frequently, as can be expected if their presumably greater featural dissimilarity to the central object does indeed attract attention.
In addition, average first fixation durations for the selected targets should be shorter when they are Central Related if priming between episodically related objects does indeed exist. Finally, if pre-attentive object processing can also cause such a priming effect, then this should become apparent in a decrease of first target fixation durations with an increase in Peripheral Relatedness, i.e. in the number of target-related objects in the periphery of the display.

3.2.2 Method

Subjects

Eight members of the University of Massachusetts subject pool participated in the experiment. All of the subjects were familiar with the eye movement registration equipment used in the experiment and none of them required corrective lenses for reading.

Stimuli

The stimuli used in this experiment were identical to those in Experiment 1 except for two minor modifications which were introduced in order to be able to reliably analyze the fixation duration data.

Specifically, because fixation durations tend to vary considerably across individuals (Rayner, 1978), it was decided to run a completely within-subjects design in order to control for the variance due to these interindividual differences. Since this necessarily implies longer experimental sessions per subject, display types 1a' and 1b' (see Table 1) were omitted in order to shorten the experiment. This had no further implications because the sole purpose of these display types was to provide intermediate levels of Peripheral Relatedness. Consequently, display types 2a' and 2b', providing extreme levels of Peripheral Relatedness, were sufficient to test the hypothesis that automatic prime processing could by itself facilitate the perception of of related objects. In addition to eliminating these two display types, the remaining display types were modified in the sense that peripheral objects no longer occupied the same position across display types. Instead, for every display type they were randomly assigned a position on one of the five corners of the imaginary pentagon. This was done in order to prevent subjects from learning to expect specific objects at specific positions, which obviously could affect both the selection of peripheral objects for fixation and the time needed to identify them.

As in Experiment 1, the six display types constructed in this manner mapped onto six context conditions. In Experiment 1, this mapping was determined by the episodic membership of the a priori designated target object. In the present experiment, however, it was determined by the episodic membership of the peripheral object that was selected for fixation by the subject. Specifically, each fixation fell into one of the six following context conditions: two conditions with a fixation-related object in the center and one or three fixation-related objects in the periphery, creating a *Central Related/Peripheral* 1 and a *Central Related/Peripheral 3* condition; two conditions with a fixation-unrelated object in the center and zero or two fixation-related objects in the periphery, producing a *Central Unrelated/Peripheral 0* and a *Central Unrelated/Peripheral 2* condition; and, finally, two conditions with no object in the center and either zero or three fixation-related objects in the periphery, resulting in a *Central Absent/Peripheral 0* and a *Central Absent/Peripheral 3* condition.

The sets of choice alternatives for the object recognition task following each display were very similar to those constructed for Experiment 1, but were modified in order to control for possible nuisance effects that could arise from the within-subjects administration of the stimuli. Specifically, rather than always probing for one of the same two targets (i.e. A1 and B1 in Table 1) following all displays constructed for a given category pair, a different target was probed for each of the six display types used in this experiment. Targets were A2 for display 1a, B2 for display

1b, B1 for display 2a, B3 for display 2b, A3 for display 2a' and A1 for display 2b'. This modification was introduced in order to avoid that subjects would learn that only a limited set of objects needed to be detected in order to perform the recognition task and consequently would attempt to detect specific object features in the displays they were presented with.

In addition, an inspection of Table 1 will reveal that this ensured that if a central object was present, there was a 50 % chance that it would be the target. Consequently, there was no reason for subjects to be biased towards neglecting the central objects, which would decrease any effect controlled prime processing might have. Furthermore, it guaranteed that if a central object was present and the target was located in the periphery, there was a 50 % chance that they would be episodically related. As a result, subjects were not induced to systematically search for either central-related or central-unrelated peripheral objects, which obviously would greatly compromise the analysis of scanning patterns across the displays. Finally, it ensured that if the target was located in the periphery, its chances of being related versus unrelated to any of the other objects in the display were 50-50. In this manner, subjects were discouraged from adopting a strategy of systematically trying to find the object that was most dissimilar to the rest of the display.

Apparatus and procedure

As in Experiment 1, stimuli, sets of choice alternatives, and mask were presented on a Hewlett-Packard 1300A CRT with a P-31 phosphor. Eye movement patterns during display exposure were monitored via a Stanford Research Institute Dual Purkinje eyetracker with a resolution of 10' of arc and a 1000 Hz sampling-rate. The eyetracker and CRT were interfaced with a Hewlett-Packard 2100 computer used for storage of the images to be presented and for keeping a complete record of saccade latencies and eye position. In this manner it was possible to terminate display exposure and present the mask once a peripheral object had received a fixation or no such fixation had occurred within a period of 1 sec following display onset.

Upon arrival, subjects received the same instructions given in Experiment 1 with the exception that they were asked to indicate the name of the object that had been in the display by tapping its serial number in the list of choice alternatives on the table in front of them. Subjects were seated 46 cm from the CRT, using a bite bar in order to eliminate head movements and keep viewing distance constant. Subsequently, the eyetracker was calibrated and subjects received two series of 56 trials, each consisting of 8 practice trials (selected from the filler stimuli constructed for Experiment 1) and 48 experimental trials. Each trial included the same events described for Experiment 1, and again no feedback was provided until the end of the experiment which lasted about 40 to 45 minutes.

Design

Each subject received all 96 experimental stimuli (i.e. 6 display types X 16 episodic category pairs) in an individually randomized order. How many observations this resulted in for each of the six context conditions naturally depended on which peripheral objects the subjects chose to fixate in the various displays. Assuming however, that subjects would choose peripheral objects at random, the distribution of the data over the context conditions should be as follows:

- 40 % of the observations for display types 1a and 1b should fall in the Central Related/Peripheral 1 condition, and 60 % in the Central Unrelated/Peripheral 2 condition.
- 80 % of the observations for display types 2a and 2b should fall in the Central Related/Peripheral 3 condition, and 20 % in the Central Unrelated/Peripheral 0 condition.
- 80 % of the observations for display types 2a' and 2b' should fall in the Central Absent/Peripheral 3 condition, and 20 % in the Central Absent/Peripheral 0 condition.

Consequently, in order to determine whether the likelihood of peripheral object selection is affected by that object's episodic relatedness to the previously fixated and/or other unattended objects in the display, one should examine how the actual proportions of the data falling into each of the context conditions deviated from what could be expected on the basis of random selection. It was therefore decided to compute the observed proportions across subjects and perform chi-square tests on the expected-observed differences for each of the three pairs of complementary proportions, i.e. i) Central Related/Peripheral 1 and Central Unrelated/Peripheral 2, ii) Central Related/Peripheral 3 and Central Unrelated/Peripheral 0, and iii) the Central Absent/Peripheral 3 and Central Absent/Peripheral 0.

As for the analysis of effects of Central and Peripheral Relatedness on first fixation durations, an analysis of variance could be performed on the generalized randomized block design (Federer, 1955; Kirk, 1982) resulting from the within-subjects administration of the context conditions. However, the inevitable presence of unequal cell n's did necessitate the use of the general linear model approach (Kirk, 1982).

3.2.3 Results

In the following analyses, all trials were excluded on which a track loss occurred (3 %), subjects failed to fixate a peripheral object within a 1 sec period following display onset (4 %), or subjects moved their eyes away from the display center in one direction and ended up fixating a peripheral object located in another direction (3 %).

The expected and observed probabilities of fixating a peripheral object in the various context conditions are presented in Table 4.

Table 4. Expected and observed probabilities of selecting a peripheral object in each of the Context conditions

CONTEXT CONDITION	EXPECTED	OBSERVED
1) Central Related/Peripheral 1	.400	.38 (n=86)
+		
2) Central Unrelated/Peripheral 2	.600	.62 (n=142)
3) Central Related/Peripheral 3	.800	.73 (n=173)
+		
4) Central Unrelated/Peripheral 0	.200	.27 (n=64)
5) Central Absent/Peripheral 3	.800	.78 (n=176)
+		
6) Central Absent/Peripheral 0	.200	.22 (n=50)

The data presented here, show that no clear discrepancies between expected and observed probabilities emerged. This was confirmed by insignificant chi-square tests of expected-observed differences in the data-distributions over the complementary conditions 1 and 2 [chi-square = .952, p > .05], conditions 3 and 4 [chi-square = 1.155, p > .05], and conditions 5 and 6 [chi-square = 1.418 p > .05].

This absence of systematic effects of Central or Peripheral Relatedness on peripheral object selection indicates that either episodic dissimilarity does not inevitably imply featural dissimilarity, or that featural dissimilarity does not determine the direction of attentional shifts in the paradigm employed in this experiment. Either way, the absence of a Central Relatedness effect runs counter to the hypothesis that the unrelated target superiority in Experiment 1 should be attributed to covert attentional shifts towards these objects. Furthermore, the failure to find any Peripheral Relatedness effects in either the Central Present or Central Absent conditions replicates the results obtained in Experiment 1. It provides converging evidence for concluding that "pop-out" effects are not a fundamental characteristic of object perception in displays containing isolated objects with different episodic category memberships. Admittedly, it could be pointed out, that in spite of the insignificant expected-observed differences in the three data-distributions there appears to be a consistent tendency towards selecting the unrelated objects (i.e. 2.2 % for the first distribution, 7% for the second, and 2.2.% for the third). In view of previous evidence for such a bias (e.g. Loftus & Mackworth, 1978) this could be taken to suggest that there is a small effect of episodic relatedness in the peripheral object selection data, which went unnoticed because of a lack of statistical power. It should be noted, however, that even if this is the case, the effect size (i.e. 2-7%) does appear too small to explain the 9-14% unrelated target superiority in Experiment 1. Also, when computed per subject and per distribution, the bias towards selecting unrelated objects did not appear to be very consistent since it was found for only 4 subjects in the first distribution, 4 in the second and 3 in the third.

The analysis of the first fixation durations revealed a significant main effect of context condition [$F(5,35) = 3.412$, $p < .05$]. Mean first fixation durations (FFD) for each of the context conditions are presented in Table 5.

Planned comparisons using the Dunn-Sidak procedure (Kirk, 1982) revealed that the 40 ms difference between conditions 1 and 2 was significant [$tDS(35) = 2.939$, $p < .05$]. This indicates that controlled object processing primes the representations of episodically related objects and facilitates subsequent perceptual processing of these objects. Additional support for this conclusion was provided by the significant 59 ms difference between conditions 3 and 4 [$tDS(35) = 3.966$, $p < .05$]. Although the latter difference was somewhat larger, suggesting an effect of the number of peripheral primes and consequently of automatic prime processing, further comparisons failed to confirm this hypothesis. Specifically, the insignificant difference between conditions 5 and 6 [*tDS*(35) = .377, *p* > .10] indicates that

Table 5. Mean first fixation durations (FFD) per Context condition (in ms)

CONTEXT CONDITION	FFD
1) Central Related/Peripheral 1	252 (n=86)
2) Central Unrelated/Peripheral 2	292 (n=142)
3) Central Related/Peripheral 3	227 (n=173)
4) Central Unrelated/Peripheral 0	286 (n=64)
5) Central Absent/Peripheral 3	245 (n=176)
6) Central Absent/Peripheral 0	251 (n=50)

automatic prime processing was insufficient to by itself facilitate the identification of primed objects. The absence of a significant difference between conditions 2 and 4 [*tDS*(35) = .394, *p* > .10] is completely in line with this finding. Finally, while the 25 ms difference between conditions 1 and 3 suggests that extrafoveal objects could perhaps exert a priming influence when presented in the company of a related foveal object, it does not reach significance [*tDS*(35) = 1.894, *p* > .10].

Admittedly, this exploration of the first fixation data for effects of central and peripheral primes could be criticized for being somewhat fragmentary in its use of multiple pairwise comparisons. It therefore appears to be worthwhile to report the results of an alternative analysis which is less susceptible to this criticism. Specifically, if the central prime effect is designated as C and the peripheral prime effect as P, the difference between the means of conditions 1 and 2 (i.e. $\delta 1$) can be regarded as reflecting C, while P is reflected by the difference between means 5 and 6 (i.e. $\delta 3$). Following an additive logic, the difference between means 3 and 4 (i.e. $\delta 2$) can then be viewed as the result of C+P. On this view, estimates of C and P can be obtained, which derive from the data of all context conditions simultaneously. Specifically, for C this estimate is provided by the equation $(\delta 1+(\delta 2-\delta 3))/2$, and for P it can be found in the equation $(\delta 3+(\delta 2-\delta 1))/2$. Following their computation for each subject, these equations yield an average C estimate of 46.5 ms and an average P estimate of 12.5 ms. Apart from confirming the conclusions of the first analysis (i.e. a significant effect of central prime processing [$t(7) = 2.83$, $p < .05$] and a non-significant effect of peripheral prime processing [$t(7) = 1.06$, $p > .10$]), it is interesting to note that the size of the central prime effect is quite comparable to that of the priming effects observed in the Henderson et al. (1987, 1988) studies (i.e. 30-60 ms).

3.2.4 General Discussion and Conclusion

The main focus of the experiments reported above, was to explore the existence and the characteristics of a priming mechanism operating between representations of objects likely to appear in the same real-world scene.

While the unrelated target superiority in the Foveal Present conditions of Experiment 1 provided only indirect evidence for the presence of such a process, direct proof was found in the decrease of first fixation durations for central-related objects in Experiment 2. Given this result, the absence in Experiment 1 of an increase in target recognition performance contingent upon the presence of either foveal or extrafoveal primes, suggested that priming does not facilitate the identification of unattended objects. Evidence in agreement with this conclusion has also been presented by Boyce, Pollatsek, and Rayner (1989) who found peripheral target detection in tachistoscopically presented scenes to be independent of the episodic relatedness between the target and other objects in the scene. These results are entirely compatible with the view (Reinitz, Wright, & Loftus, 1989; Warren & Morton, 1982) that meaning-based priming does not facilitate object identification by bringing a conceptual object representation so close to the activation level required to set off an identification response that a minimal amount of consistent visual information is sufficient to pass this threshold. Rather, priming aids the identification of isolated objects by increasing the rate at which the object's features are extracted from the image in a capacity-limited process (Biederman, Blickle, Teitelbaum, & Klatsky, 1988) requiring both visual acuity and a sufficiently selective allocation of visual attention (Strong & Whitehead, 1989).

In addition to this need for attention to the target in order for priming to surface, controlled processing of the prime seems to be a necessary condition for priming to originate, as indicated by the absence of a significant effect of peripheral primes on first fixation durations in Experiment 2. While this conclusion is corroborated by the Henderson et al. (1988) failure to find a target-display consistency effect, it could perhaps be viewed as too strong given the suggestion of a peripheral prime effect in Experiment 2 conditional upon the simultaneous presence of a foveal prime. However, contrary to the Henderson et al. experiments, the task conditions in this experiment did not favor a standard sequence for leisurely fixating all objects in the displays. As a result it is not impossible that an initially wide distribution or even a selective covert shift of attention may have preceded the first peripheral object fixation, thus allowing for a facilitation of peripheral prime processing by the central prime. Such a facilitation could then lead to sufficiently detailed pattern recognition of peripheral primes, in order to cause an additional priming benefit for the first related object that is fixated. Consequently, while this would perhaps challenge a strict sequential model of visual attention (Henderson et al., 1989), it would still be consistent with the claims that priming requires the allocation of attention to the prime and will only facilitate extrafoveal information use for attended objects (Henderson et al., 1988).

Based on the above characterization of episodic priming, some predictions can be made which should allow for an unambiguous test of the role of this mechanism in the appearance of object probability effects on object identification in scenes. Specifically, if episodic inter-object priming is solely responsible for these effects, object identification should not be easier for 1) probable objects that are the first object attended to in a scene, 2) probable objects that are not selectively attended to, and 3) probable objects attended to after attending to an improbable object.

As was already mentioned, research reported by Biederman and colleagues clearly contradicts all three predictions in its claim to have established superior perceptibility for any probable object at any position in any natural scene that is presented for a mere 100-150 ms. Recently, this claim has been reaffirmed in a study (Boyce, Pollatsek, & Rayner, 1989) revealing that target-background consistency is a crucial determinant of this immediate object probability effect on pre-cued target detection, while the target's episodic relatedness to other objects in the display is wholly irrelevant. However, as argued in Chapter 1, this kind of research can be criticized for its susceptibility to the influence of post-perceptual response strategies obscuring genuine perceptual effects of scene context on object processing. The large reduction in task performance when, as in Experiment 1, accuracy of forced choice recognition is measured rather than accuracy of simple present-absent decisions, only serves to strengthen this suspicion.

Moreover, even if it can be assumed that the observed effects are indeed perceptual, they are by no means mandatory, as some authors claim (Klatsky et al., 1981). Recently, De Graef, Christiaens and d'Ydewalle (1990) reported on the context-sensitivity of first fixation durations for objects incidentally fixated during the free exploration of a scene in search for non-objects (Kroll & Potter, 1984). With respect to object probability effects, they found that such effects clearly did not surface from the very first scene fixation on but rather developed gradually over the course of scene exploration. At the very least, this leaves open the possibility that, under some conditions, individual object processing and subsequent inter-object priming may be at the basis of object probability effects. Specifically, when one looks at the differences between a task allowing only one glimpse at a scene in order to detect an object at an uncertain position and a task allowing for free scene exploration in search for non-objects, the crucial determinant of differences in the characteristics of context effects may be the mode of attentional distribution adopted by the viewer. Indeed, the former type of task is much more likely to favor a wide distribution of attention across the entire scene, favoring the extraction and use of low-resolution, global information such as scene background. In the latter type of task, however, a sequential distribution of attention to successively fixated potential objects is more appropriate, favoring the operation of mechanisms such as the episodic priming described above. Consequently, unless the latter task can be demonstrated to be a less adequate approximation of everyday scene perception, a rejection of inter-object priming as an account of object probability effects in scenes will have to based on a direct test of the three forementioned predictions under task conditions that do not discourage or even prevent selective attention to individual objects in the scene. Until then, an identification advantage for the probable objects in a real-world scene can not be taken as a sound basis for rejecting data-driven models of object identification in favor of concept-driven accounts.

APPENDIX A: EXPERIMENTAL STIMULI

EPISODICAL CATEGORIES	EXTRAFOVEAL TARGETS	FOVEAL PRIMES	EXTRAFOVEAL PRIMES		
1) Gas Station	car	gaspump	motor-cycle	jerry can	bus
Orchestra	piano	trumpet	violin	flute	horn
2) Laundry Room	iron	shirt	dress	tie	vest
Playgrounds	football helmet	football	baseball bat	tennis-racket	baseball glove
3) Farm	horse	sheep	rooster	scythe	cat
Living Room	rocking chair	televi-sion	dresser	lamp	vacuum cleaner
4) Street	bus	fire hydrant	bicycle	traffic light	motor-cycle
Kitchen	stove	fridge	scale	egg timer	garbage can
5) Dinner Table	fork	salt shaker	wine-glass	bottle	pot
Toolshed	axe	hammer	paint-brush	saw	file
6) Garden	watering can	bucket	shears	spade	hose
Laundry Room	sweater	pants	ironing board	dress	tie
7) Pole	penguin	igloo	seal	polar bear	walrus
Farm	pig	barn	sheep	scythe	tractor
8) Kitchen	kettle	frying pan	rolling pin	pitcher	blender
Bathroom	blowdryer	comb	tooth-paste	tooth-brush	hair-brush

APPENDIX A (Continued)

9) Playgrounds	wagon	tricycle	roller skates	kite	skate board
Office	telephone	desk	computer	stapler	envelope
10) Street	traffic light	bicycle	parking meter	bus	ambu- lance
Farm	cow	chicken	goose	pig	tractor
11) Garden	roller	pitch- fork	wheel- barrow	spade	rake
Living room	couch	table	vacuum cleaner	chair	lamp
12) Toolshed	drill	pliers	paint- brush	screw- driver	wrench
Vegetable Stand	pepper	carrot	arti- choke	onion	mushroom
13) Fruit Basket	apple	banana	pear	pine- apple	melon
Playgrounds	skate board	roller skates	top	tri- cycle	baseball bat
14) Bedroom	dresser	bed	alarm- clock	chair	lamp
Street	motorcycle	parking meter	car	truck	ambu- lance
15) Farm	goat	rooster	horse	milkcan	dog
Orchestra	drum	guitar	trumpet	xylo- phone	flute
16) Garden	wheelbarrow	spade	roller	hose	shears
Kitchen	coffee pot	funnel	toaster	blender	rolling pin

APPENDIX B: CHOICE SETS FOR EXPERIMENTAL STIMULI

EPISODICAL CATEGORIES	TARGETS	DISTRACTORS		
1) Gas station	car	airpump	fire hydrant	truck
Orchestra	piano	drum	xylophone	saxophone
2) Laundry room	iron	pants	ironing board	skirt
Playgrounds	football helmet	frisbee	skateboard	ball
3) Farm	horse	goose	barn	dog
Living room	rocking chair	turntable	couch	vase
4) Street	bus	car	ambulance	parking meter
Kitchen	stove	frying pan	pot	blender
5) Dinner table	fork	spoon	pitcher	ladle
Toolshed	axe	drill	wrench	pliers
6) Garden	watering can	rake	roller	lawnmower
Laundry room	sweater	clothes rack	iron	skirt
7) Pole	penguin	whale	harpoon	sled
Farm	pig	horse	dog	well
8) Kitchen	kettle	pot	toaster	egg timer
Bathroom	blowdryer	soap	razor	shaving brush
9) Playgrounds	wagon	tennis-racket	baseball bat	ball
Office	telephone	paperpunch	ashtray	typewriter

APPENDIX B (Continued)

10) Street	traffic light	truck	mailbox	motorcycle
Farm	cow	dog	sheep	barn
11) Garden	roller	hose	lawnmower	spray can
Living room	couch	vase	telephone	dresser
12) Toolshed	drill	file	saw	hammer
Vegetable stand	pepper	corn	tomato	cellery
13) Fruit basket	apple	coconut	orange	grapes
Playgrounds	skateboard	slingshot	wagon	scooter
14) Bedroom	dresser	pajamas	radio	pillow
Street	motorcycle	fire hy-drant	bus	mailbox
15) Farm	goat	sheep	cow	tractor
Orchestra	drum	horn	piano	violin
16) Garden	wheelbarrow	rake	lawnmower	lawnchair
Kitchen	coffee pot	grill	cup	pot

APPENDIX C: PRACTICE AND CONTROL STIMULI

FOVEAL TARGETS	EXTRAFOVEAL UNRELATED OBJECTS				
1) shoe	door	broom	knife	tie	squirrel
2) mouse	whistle	light-bulb	glass	moon	boat
3) anchor	banana	leaf	onion	bell	screw
4) cannon	funnel	key	pipe	window	pitcher
5) sock	cup	broom	flute	apple	star
6) whistle	squirrel	pipe	baseball bat	pear	bell
7) spoon	tooth-paste	moon	bat	rabbit	anchor
8) lock	pineapple	light-bulb	leaf	hair-brush	tie
9) flower	shoe	key	glass	cat	window
10) hat	egg timer	cannon	door	banana	stool
11) pitcher	fire hy-drant	onion	boat	scythe	screw
12) broom	mouse	tooth-brush	rolling pin	chicken	knife
13) scissors	bell	vacuum cleaner	jerry can	goat	envelope
14) pipe	pear	lamp	comb	file	mushroom
15) glass	sock	whistle	lock	flower	hat
16) squirrel	tooth-paste	cup	pine-apple	shoe	egg timer
17) lightbulb	broom	pipe	moon	key	cannon
18) bat	flute	roller skate	leaf	glass	door

APPENDIX C (Continued)

19) rabbit	apple	paint-brush	cat	tie	baseball bat
20) star	bell	anchor	banana	window	stool
21) chicken	pitcher	broom	scissors	pipe	mushroom
22) envelope	ambulance	mouse	whistle	pear	leaf
23) knife	onion	tooth-brush	vacuum cleaner	jerry-can	file
24) screw	boat	rolling pin	lamp	comb	goat

APPENDIX D: CHOICE SETS FOR PRACTICE AND CONTROL STIMULI

TARGETS	DISTRACTORS		
1) shoe	boot	skate	clog
2) mouse	cheese	cat	rat
3) anchor	surfboard	boat	raft
4) cannon	tank	bazooka	machine gun
5) sock	shoe	pants	boot
6) whistle	horn	flute	rattle
7) spoon	knife	pitcher	rolling pin
8) lock	key	chain	door
9) flower	bee	tree	leaf
10) hat	coat rack	scarf	umbrella
11) pitcher	funnel	cup	knife
12) broom	dustpan	mop	bucket
13) scissors	needle	thread	button
14) pipe	ashtray	cigarette	lighter
15) glass	bottle	cup	jug
16) squirrel	tree	nut	cat
17) lightbulb	socket	lightswitch	lamp
18) bat	owl	mouse	cat
19) rabbit	hedgehog	fox	squirrel
20) star	moon	sattelite	rocket
21) chicken	egg	coop	chick
22) envelope	paper clip	pen	stamp
23) knife	bread	butter	cheese
24) screw	nail	hammer	file

NOTES

1. Naturally, a failure to observe any target recognition accuracy differences in the present experiment would be difficult to interpret as showing that priming is indeed constrained to controlled object-processing. Since prime-target relatedness in all stimuli is defined episodically rather than semantically, the absence of any effects could also mean that priming is constrained to semantically related objects. It was decided to filter out this interpretational problem in a separate experiment should it arise.

BIBLIOGRAPHY

Antes, J.R., Mann, S. M., & Penland, J. G. (1981). Local precedence in picture naming: The importance of obligatory objects. Paper presented at the 1981 meeting of the Psychonomic Society.

Antes, J. R., & Penland, J. G. (1981). Picture context effects on eye movement patterns. In D. F. Fischer, R. A. Monty & J. W. Senders (Eds.), Eye movements: Cognition and visual perception. Hillsdale, N. J.: Erlbaum.

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

Antes, J. R., Singsaas, P. A., & Metzger, R. L. (1978). Components of pictorial informativeness. Perceptual and Motor Skills, 47, 459-464.

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

Biederman, I. (1982). Human information processing of real-world scenes (Final Report on U.S. Army Research Institute Grant MDA 903-79-G). Buffalo: State University of New York at Buffalo, Department of Psychology.

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

Biederman, I. (1988). Aspects and extensions of a theory of human image understanding. In Z. W. Pylyshyn (Ed.), Computational processes in human vision: An inter-disciplinary approach (pp. 370-428). Norwood, NJ: Ablex.

Biederman, I., Blickle, T., Teitelbaum, C., & Klarsky, G. (1988). Object search in nonscene displays. Journal of Experimental Psychology: Learning, Memory, and Cognition, 14, 456-467.

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

Bobrow, D. G., & Norman, D. A. (1975). Some principles of memory schemata. In D. G. Bobrow & A. Collins (Eds.), Representation and understanding: Studies in cognitive science. New York: Academic Press.

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

Breitmeyer, B. G., & Ganz, L. (1976). Implications of sustained and transient channels for theories of visual pattern masking, saccadic suppression and information processing. Psychological Review, 83, 1-36.

Carr, T., McCauley, C., Sperber, R., & Parmelee, C. (1982). Words, pictures and priming: On semantic activation, conscious identification, and the automaticity of information processing. Journal of Experimental Psychology: Human Perception and Performance, 8, 757-777.

De Graef, P., Christiaens, D., & d'Ydewalle, G. (1990). Perceptual effects of scene context on object identification. To appear in: Psychological Research.

Egeth, H., Jonides, J., & Wall, S. (1972). Parallel processing of multielement displays. Cognitive Psychology, 3, 674-698.

Federer, W. (1955). Experimental design: Theory and application. New York: Macmillan.

Fischler, M. A. (1978). On the representation of natural scenes. In A. R. Hanson & E. M. Riseman (Eds.), Computer vision systems. Orlando, FLA: Academic Press, Inc.

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

Guzman, A. (1969). Decomposition of the visual field into three-dimensional bodies. In A. Grasselli (Ed.), Automatic interpretation and classification of images. New York: Academic Press.

Hanson, A. R., & Riseman, E. M. (1978). VISIONS: A computer system for interpreting scenes. In A. R. Hanson & E. M. Riseman (Eds.), Computer vision systems. Orlando, FLA: Academic Press, Inc.

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. (1988). Extrafoveal information use during object identification. Unpublished manuscript.

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

Hoffman, D., & Richards, W. (1985). Parts of recognition. Cognition, 18, 65-96.

Huttenlocher, J., & Kubicek, L. F. (1983). The source of relatedness effects on naming latency. Journal of Experimental Psychology: Learning, Memory and Cognition, 9, 486-496.

Ikeda, M., & Takeuchi, T. (1975). Influence of foveal load on the functional visual field. Perception & Psychophysics, 18, 255-260.

Kirk, R. E. (1982). Experimental design: Procedures for the Behavioral Sciences. Monterey, CA: Brooks/Cole.

Klatsky, G. J., Teitelbaum, R. C., Mezzanotte, R. J., & Biederman, I. (1981). Mandatory processing of the background in the detection of objects in scenes. Proceedings of the Human Factors Society, 25, 272-276.

Kosslyn, S. M. (1987). Seeing and imaging in the cerebral hemispheres: A computational approach. Psychological Review, 94, 148-175.

Kroll, J. F., & Potter, M. C. (1984). Recognizing words, pictures and concepts: A comparison of lexical, object and reality decisions. Journal of Verbal Learning and Verbal Behavior, 23, 39-66.

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.

Marr, D. (1978). Representing visual information: A computational approach. In A. R. Hanson & E. M. Riseman (Eds.), Computer vision systems. Orlando, FLA: Academic Press, Inc.

Marr, D. (1982). Vision: A computational investigation into the human representation and processing of visual information. San Francisco: W. H. Freeman & Company.

McArthur, D. J. (1982). Computer vision and perceptual psychology. Psychological Bulletin, 92, 283-309.

McCauley, C., Parmelee, C., Sperber, R., & Carr, T. (1980). Early extraction of meaning from pictures and its relation to conscious indentification. Journal of Experimental Psychology: Human Perception and Performance, 6, 265-276.

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

Minsky, M. (1975). A framework for representing knowledge. In P. H. Winston (Ed.), The psychology of computer vision. New York: McGraw Hill.

Morrison, R. E. (1984). Manipulation of stimulus onset delay in reading: Evidence for parallel programming of saccades. Journal of Experimental Psychology: Human Perception & Performance, 10, 667-682.

Paap, K, & Newsome, S. (1981). Parafoveal information is not sufficient to produce semantic or visual priming. Perception & Psychophysics, 29, 457-466.

Palmer, S. E. (1975). Visual perception and world knowledge: Notes on a model of sensory-cognitive interaction. In D. A. Norman & D. E. Rumelhart (Eds.), Explorations in cognition. San Francisco: W. H. Freeman & Company.

Rayner, K. (1978). Eye movements in reading and information processing. Psychological Bulletin, 85, 618-660.

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

Reinitz, M., Wright, E., & Loftus, G. (1989). Effects of semantic priming on visual encoding of pictures. Journal of Experimental Psychology: General, 118, 280-297.

Schank, R. C., & Abelson, R. (1977). Scripts, plans, goals and understanding. Hillsdale, N. J.: Erlbaum.

Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing: 2. Perceptual learning, automatic attending, and a general theory. Psychological Review, 84, 127-190.

Snodgrass, J. G., & Vanderwart, M. (1980). A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity. Journal of Experimental Psychology: Human Learning and Memory, 6, 174-215.

Sperber, R., McCauley, C., Ragain, R., & Weil, C. (1979). Semantic priming effects on picture and word processing. Memory & Cognition, 7, 339-345.

Strong, G., & Whitehead, B. (1989). A solution to the tag-assignment problem for neural networks. Behavioral and Brain Sciences, 12, 381-433.

Ungerleider, L. G., & Mishkin, M. (1982). Two cortical visual systems. In D. J. Ingle, M. A. Goodale & R. J. W. Mansfield (Eds.), Analysis of visual behavior. Cambridge, MA: MIT Press.

Waltz, D. (1975). Understanding line drawings of scenes with shadows. In P. H. Winston (Ed.), The psychology of computer vision. New York: McGraw Hill.

Warren, C., & Morton, J. (1982). The effects of priming on picture recognition. British Journal of Psychology, 73, 117-129.

Wheeler, D. D. (1970). Processes in word recognition. Cognitive Psychology, 1, 59-85.