

This is to certify that the
dissertation entitled

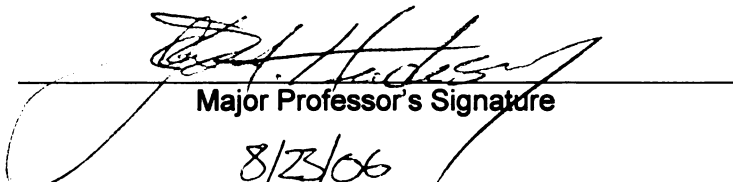
THE INTEGRATION OF INFORMATION ABOUT OBJECTS
ACROSS EYE MOVEMENTS

presented by

DANIEL A. GAJEWSKI

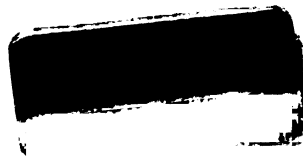
has been accepted towards fulfillment
of the requirements for the

Doctoral degree in Psychology


Major Professor's Signature

8/23/06

Date



LIBRARY
Michigan State
University

PLACE IN RETURN BOX to remove this checkout from your record.
TO AVOID FINES return on or before date due.
MAY BE RECALLED with earlier due date if requested.

DATE DUE	DATE DUE	DATE DUE

**THE INTEGRATION OF INFORMATION ABOUT OBJECTS
ACROSS EYE MOVEMENTS**

By

Daniel A. Gajewski

A DISSERTATION

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

DOCTOR OF PHILOSOPHY

2006

ABSTRACT

THE INTEGRATION OF INFORMATION ABOUT OBJECTS ACROSS EYE MOVEMENTS

By

Daniel A. Gajewski

This work investigated the nature of the information about objects that is maintained and integrated across eye movements. Given that eye movements are needed to bring objects from the periphery to the fovea so that visual details can be resolved, what information acquired about an object in the periphery before the saccade plays a functional role in the identification of that object when it is fixated? In the extrafoveal preview paradigm, participants direct their eyes to a peripherally presented preview object that is replaced during the saccade with a to-be-named target object. Inferences about the nature of the information integrated are made on the basis of preview benefits, which are the differences in naming latencies when the preview and target objects are similar versus dissimilar.

The current study determined the role for visual integration in the generation of preview benefits using full-color pictures of real-world objects and a non-repeating stimulus set. Preview and target objects were from the same basic-level category but varied in terms of visual similarity. Preview benefits were observed for identical and a range of visually dissimilar previews compared to meaningless-object and different-object controls. These effects were observed despite the fact that items were not repeated. The magnitudes of the preview benefits were largely undiminished by surface-level differences between the preview and target, such as color, texture, and the shape of the parts, suggesting that the representations involved are abstracted away from these

properties. Preview benefits were reduced, however, by differences in viewpoint generated by rotation, mirror-reversal, or by taking an entirely different perspective of the object.

To determine the extent to which facilitation depends on identification of the preview object before the saccade, a regression analysis was employed examining preview benefits as a function of the proportion of participants who correctly identified each object on the basis of a peripheral preview alone. Preview benefits increased as objects were more readily identified from the periphery, and the better part of the effect depended on identification accuracy; however, preview benefits were also observed for objects that never or rarely were identified from the periphery.

A second regression analysis examined the combined roles for visual similarity and preview identification. Here, the importance of maintaining an object's viewpoint across saccades was confirmed, but the effect of viewpoint-similarity did not depend on the identifiability of the preview object. The pattern of results suggests that preview benefits are enhanced by but not dependent on identification before the saccade, and that integration at the level of object identity combines additively with that provided by the integration of visual information.

ACKNOWLEDGEMENTS

I would like to thank John Henderson, Rose Zacks, Erik Altman, and Fred Dyer for their helpful comments on the dissertation. I also wish to thank Gary Schrock and Christy Miscisin for their technical assistance as well as a number of research assistants who have assisted with various aspects of this project, ranging from stimulus generation to the collection of data: Jennifer Gorman, Lyaz Marshall, Paula Ogston, Tonisha Banks, Matt Piszczek, and Twila Starosciak.

Portions of this research were supported by NSF-IGERT Grant ECS-9874541. This material is based upon work supported by, or in part by, the U. S. Army Research Office under grant number W911NF-04-1-0078 awarded to John Henderson.

TABLE OF CONTENTS

List of Tables	vi
List of Figures	vii
Introduction	1
The Spatiotopic Fusion Hypothesis	1
The Continuation of Processing Across Saccades	3
Transsaccadic Object Identification	7
The Role of Spatial Location	9
The Two Representational Systems Theory	14
Overview of Current Research	16
Experiment 1	20
Methods	23
Results and Discussion	27
Experiment 2	30
Methods	33
Results and Discussion	34
Experiment 3	37
Methods	42
Results and Discussion	45
General Discussion	55
Appendix A	66
Appendix B	69
Appendix C	71
Appendix D	72
Appendix E	74
Appendix F	77
References	79

LIST OF TABLES

Table 1. Mean Naming Latencies and Standard Errors (in milliseconds) for Experiments 1 and 2 by Preview Condition.	28
Table 2. Correlations for the Measures Employed in Experiment 3 (* $p < .01$)	49
Table 3. Beta Weights From Regression Analysis I (* $p < .05$).	50
Table 4. Correlations of Preview Benefits with Visual Similarity Ratings (* $p < .01$).	52
Table 5. Beta Weights From Regression Analysis II (* $p < .05$).	54

LIST OF FIGURES

Figure 1. Schematic illustration of the displays presented in Experiment 1 (Top). During the first display, participants fixated a small plus sign on the left-hand side of the screen. A preview object was then presented in the second display and participants initiated a saccade toward the object. While the eyes were moving, the display was changed to present the target object. Participants named the target object as quickly as possible after the saccade. Example items for the meaningless-object, different-object, and identical preview conditions are also shown (Bottom). Full-color images were used in the actual experiment. The trial illustration is not shown to scale.	21
Figure 2. Example stimuli from Experiment 2. The columns from left to right correspond to the identical, visually-dissimilar, and maximally-dissimilar preview conditions. The objects in the identical and visually-dissimilar conditions were the same as those employed in Experiment 1. Most objects in the maximally-dissimilar condition were different exemplars from an absolutely different perspective (the camera), some were different exemplars mirror-reversed (the pen), and a few were substantially different in another way (the apple). Full-color images were used in the actual experiment.	32
Figure 3. Frequency distribution for the proportions of participants who correctly identified a given object from the periphery in the extrafoveal identification task.	47
Figure 4. Preview benefits (in milliseconds) as a function of extrafoveal identification	48

INTRODUCTION

Human vision is dynamic: one's perception of a scene is a product of a sequential sampling process. Because the foveal region of high visual acuity covers an area corresponding to only about two degrees of visual angle, the eyes are directed from one point to another at a rate of nearly three times per second to resolve and encode the details. In addition, information is extracted from the environment primarily during fixation, when the point of regard is stable (Matin, 1974; Rayner, 1998). As a result, the perception of a scene can be characterized as a series of relatively discrete glimpses or snapshots of the world, with high-resolution information available only at the center of vision. The dynamic properties of visual information acquisition, coupled with the variable resolution of the input, lead to a number of empirical questions. The current research is concerned with the nature of the representations that are integrated from one fixation to the next. In what sense does information acquired from one fixation carry over and combine with information acquired during the subsequent fixation? Must visual processing begin anew with each fixation?

The Spatiotopic Fusion Hypothesis

The initial hypothesis about the combining of information across saccades was most literal. The idea, labeled the *spatiotopic fusion hypothesis* by Irwin (1993), was that contents of individual fixations could be melded together within a spatiotopic buffer to form a stable, coherent percept. The fusion hypothesis was stated formally by a number of theorists (e.g., Feldman, 1985; McConkie & Rayner, 1976; Trehub, 1977), but the idea is actually quite old. In cognitive psychology's original text, Neisser (1967) suggested that it was common to assume that successive snapshots are projected "onto the right

place in a higher-level ‘map’ of phenomenal space” (p. 140), and by most accounts the idea can be traced back to Helmholtz ([1867] 1925).

The idea of successive glimpses translated and fused together within a spatiotopic reference frame is appealing because it could simultaneously account for a number of puzzles of human visual perception. By maintaining information from previous glimpses in a spatially organized buffer, the spatiotopic fusion hypothesis gives memory a prominent role in one’s immediate perception of the world, potentially explaining why disruptions associated with eye movements go unnoticed, why the percept seems to provide more detail than is available in the retinal input at any given point in time, and how the positions of objects in the world with respect to the viewer are perceived as stable despite the retinal flux, a phenomenon called *visual direction constancy*. However, while there was some initial support for the idea (e.g., Breitmeyer, Kropfl, & Julesz, 1982; Davidson, Fox, & Dick, 1973, Jonides, Irwin, & Yantis, 1982; Ritter, 1976; Wolf, Hauske, & Lupp, 1978; 1980), the early findings have been countered by a number of negative results. One informative task involves the presentation of two arrays of dots, one before and one after a saccade. The arrays are created by randomly placing 24 dots in a 5 by 5 matrix so that one set of 12 appear in a presaccadic array and a different set of 12 appear in a postsaccadic array. Participants are required to report the single location in the matrix that is left unfilled. Because the arrays appear at the same location in space but at different retinal locations, successful report of the missing dot should occur only if the two arrays are fused spatiotopically. Early successes in tasks of this nature (e.g., Breitmeyer et al., 1982; Jonides et al., 1982) could not be replicated and were attributed to phosphor persistence (Irwin, Yantis, & Jonides, 1983; Rayner & Pollatsek, 1983).

Although the decay rate of the screen phosphor in the original experiments was thought to be sufficiently fast, careful examination of the issue suggested that there was enough residual illumination from the initial array of dots that the two arrays could to some degree be seen at the same time.

Performance failures in the dot matrix task were echoed in tasks that required the combining of line of segments to form three-letter words (O'Regan & Lévy-Schoen, 1983), the summation of sine wave gratings to enhance spatial frequency judgments (Irwin, Zacks, & Brown, 1990), and the spatiotopic combining of a mask or bar probe with the location of a letter within an array (Irwin, Brown, & Sun, 1988). The results of an early study employing an alternating letter case paradigm were also inconsistent with the spatiotopic fusion hypothesis. McConkie and Zola (1979) had participants read passages of text with words composed of alternating upper and lower case letters (e.g., ThE fLoRiDa EvErGlAdEs). During the reading of the text a number of saccades were selected on the basis of an eye velocity criterion, and during these saccades the text was either changed so that every letter switched case (e.g., tHe FlOrIdA eVeRgLaDeS) or the text remained unchanged. Fixation durations, saccade lengths, and regressive saccades were not affected by this manipulation; participants did not even notice that these changes were taking place while they were reading. If pattern information were spatiotopically aligned and combined across saccades, these changes should have been salient and reading should have been disrupted.

The Continuation of Processing Across Saccades

While the spatiotopic fusion hypothesis does not appear to be a valid conceptualization for transsaccadic integration, investigations of the perceptual span in

reading supported the idea that information is integrated across eye movements in one way or another. For example, in an early study using a saccade-contingent boundary technique, Rayner (1975) had participants read short passages that contained a number of critical words that were sometimes altered until the gaze position crossed a software-defined boundary. Fixation durations after the boundary crossing were shorter when the word before the crossing shared properties of the word after the crossing, such as beginning letters or word shape, suggesting that information acquired during the pre-crossing fixation contributes in some way to processing during the post-crossing fixation.

To investigate this process more directly, Rayner (1978; Rayner, McConkie, & Ehrlich, 1978) developed an extrafoveal preview paradigm. In this paradigm, participants initiate a saccade toward a location in the periphery where a preview word is presented. While the eyes travel toward the preview item (a word or nonword), a saccade-contingent display change is executed and the preview item is replaced with a to-be-named target word. The underlying assumption is that identification and naming will be facilitated to the extent that information provided by the preview is retained and integrated with information acquired when the target is fixated. The kind of information that is integrated can be explored within this paradigm by manipulating the similarity between the preview and target items. Preview benefits, the difference in naming latencies when the preview and target items are similar versus dissimilar, should be observed when the additional information provided by the similar preview is maintained and integrated. In their studies, the orthographic similarity of preview and target items was manipulated. Target words (e.g., phone) were named fastest when the preview was identical. However, naming was faster when the preview was a word (plane) or a nonword (ptcne) that shared the target's

initial and final letters and maintained the target's shape than when only the terminal letters (psfne) or only the shape (qtcuc) were maintained. These data suggested that information about terminal letters and word shape is maintained and integrated across eye movements but not lexical or semantic information, which bolstered the argument that extrafoveal information acquired during one fixation could be maintained and integrated across an eye movement, and that this process supports the identification of words when they are directly fixated.

The properties of transsaccadic integration in the context of reading and word recognition have subsequently been elaborated by Rayner and colleagues (Balota & Rayner, 1983; Pollatsek, Lesch, Morris, & Rayner, 1992; Rayner, McConkie, & Ehrlich, 1978; Rayner, McConkie, & Zola, 1980), primarily using the extrafoveal preview paradigm. Rayner et al. (1978), for example, demonstrated that the benefit associated with the extrafoveal preview is not tied to the execution of the saccade, but when an eye movement is executed, facilitation occurs only for the region toward which the eyes are moving. This study also provided evidence that the effect was one of facilitation as opposed to interference. All alternate word preview conditions showed a benefit relative to a no-extrafoveal-stimulus condition (a single asterisks). Rayner et al. (1980) further investigated the kind of overlap between the preview and target words that was necessary to produce a benefit. Preview benefits were observed when the preview and target shared two or three initial letters (green-grave or grain-grave), but these benefits were not as large as when the preview and target were identical. There was no benefit associated with previews that shared only the first letter (write-walks), that shared all four ending letters (write-trite), that was semantically related (write-print), or that shared the first phoneme

(write-rough). Pollatsek et al. (1992), however, did show a phonological effect in this task. Preview benefits for homophones were somewhat greater than for pairs matched in terms of visual similarity (cite-site versus cake-sake).

The view of transsaccadic integration that has emerged from these studies is one of continued processing of text based on letter codes abstracted away from precise visual form (e.g., type style and letter case). Words are processed in the periphery, but this processing does not proceed to the point of complete identification. Rayner et al. (1978) argue that if the preview were fully identified before the saccade, naming would suffer from interference when the preview was a word other than the target. However, their data showed facilitative effect even when this was this case. In addition, fully identified preview words that differ from the target might be expected to intrude during naming, which did not occur. The integration process, then, according to Pollatsek et al. (1992), can be characterized in one of two ways, depending on one's "modeling taste" (p. 159). First, integration could be explained in terms of the activation of abstract letter codes (graphemes) as well as phonemes. These orthographic and phonologic units remain active across the saccade, which shortens the time needed for identification once fixated. Second, integration could be explained in terms of activation of a neighborhood of entries in the lexicon. Because neighborhood activation is thought to be influenced by its similarity to the information coded from the preview, the latter of these two might be better suited to deal with the fact that identical words provide more facilitation than do homophones or words that share the first three letters, which suggests that factors like word shape also come into play.

The continued-processing view of transsaccadic integration has done equally well in the context of the viewing of pictorial stimuli. In an initial study of transsaccadic integration for real-world objects, Pollatsek, Rayner, and Collins (1984) manipulated the visual, conceptual, and name similarity of preview and target objects using line drawings in the extrafoveal preview paradigm. In their study, target objects were named more quickly when the preview and target objects were identical. However, the amount of facilitation was not reliably diminished when the size of the object changed from one fixation to the next. In addition, preview benefits were observed when the preview and target objects were different exemplars from the same basic-level category. Although the results suggested that veridical representations are not integrated across eye movements (see also Henderson, 1997), additional experiments supported the idea that there is a visual component to the effect over and above that derived from the objects belonging to the same category or having the same name: target objects (e.g., a ball) were named faster after previews that were visually similar (e.g., a tomato) than after those that were semantically similar (e.g., a bat), and greater facilitation was observed when the preview was a mirror image of the target object than when it was a different object with the same name (e.g., a baseball bat and a flying-mammal bat). On the basis of this study, Pollatsek et al. concluded that integration occurs at the level of the visual features of the object as well as its name.

Transsaccadic Object Identification

The continued-processing framework applied to pictorial stimuli returns transsaccadic integration to the domain of scene perception, but now information integration is posited to play a role in the identification of objects as opposed to the compiling of a highly-

detailed internal picture of the world. That is, given that the perception of a scene largely entails the sequential fixation of objects, and that eye movements serve to bring objects from the peripheral to the central region of the visual field so that the details can be resolved, what information acquired about an object from beyond fixation has a functional role in the identification of that object when it is fixated?

The problem is best understood when the relationship between eye movements and attention is considered. The most prevalent view of the saccade-attention dynamic is one where a shift of attention to the location of the upcoming saccade target precedes the change in gaze direction (Henderson, 1992b; Henderson, Pollatsek, & Rayner, 1989; Hoffman & Subramaniam, 1995; Kowler, Anderson, Doshier, & Blaser, 1995; Shepherd, Findlay, & Hockey, 1986). Interestingly, preview benefits were used as a tool to investigate the allocation of attention in the context of eye movements in much the same way that they were used to investigate the perceptual span in reading. Henderson et al. (1989) had participants sequentially view an array of four objects arranged in a square in preparation for a memory test. The availability of the objects during the course of viewing was manipulated using a moving window technique. Fixation durations on the objects were shorter when the full display was available during the entire viewing sequence than when the objects were presented one at a time as they were foveated. Objects were also fixated more briefly in a condition where two objects were presented at a time, the foveated object and the next object in the viewing sequence. Importantly, the full-display condition did not provide an additional advantage over the foveated plus next object condition, suggesting that extrafoveal information acquisition was limited to the object that was about to be fixated.

Sequential attention models such as the one proposed by Henderson (1992b) suggest that attention is initially allocated to the foveated stimulus. When processing at the center of fixation is complete or nearly complete, attention is disengaged and reallocated to a more peripheral location. This reallocation of attention coincides with the programming of an eye movement that brings the center of vision to the newly attended region of the visual field. Importantly, the lag between the shift of attention and the execution of the eye movement affords the visual system a blurry glimpse of the object that is the target of the impending saccade. Transsaccadic integration can thus be thought of as a combining of processing initiated on the object at the extrafoveal region of the retina with that initiated when the object is foveated.

The Role of Spatial Location

A question that has been raised about the integration of information about objects is whether facilitation of processing depends on the object occurring in the same location before and after the saccade. The question is important because the location dependence or independence of the effect provides information regarding the kind of representational systems that play a role in transsaccadic integration. The candidates will be referred to here as the *object type* and *object token* representational systems. The object type system is responsible for the identification of objects. Location-independent benefits are generally taken to suggest the priming of long-term memory representations stored within this system. These would correspond to the visual descriptions that support object identification as well as conceptual identity and name codes. This source of facilitation would not be expected to depend on the location of the object in the preview display because the system that supports object identification is generally thought to be

independent of the system that supports object localization, as suggested by dissociable effects of damage to what has become known as the *what* and *where* pathways (Ungerleider & Mishkin, 1982).

The object token system, on the other hand, is responsible for maintaining information about objects as they move or change (Kanwisher & Driver, 1992; Kahneman, Treisman, & Gibbs, 1992). An influential theoretical instantiation of a token system involves a construct termed the *object file* (Kahneman & Treisman, 1984). Object files are short-term, episodic representations. Importantly for the present discussion, they are thought to be addressed by spatial and temporal coordinates rather than by form or identity (Kahneman et al., 1992, Kanwisher & Driver, 1992, Treisman, 1993). The construct is founded on the idea that an object is an object by way of its continuity in space and time, a concept that is often illustrated with an example of apparent motion. Consider a movie of a simple object, such as a square, translating across a computer display so that when it reaches mid-screen it is replaced with a triangle. The perception of motion, of course, is created by controlling the displacement of the object from one frame to another in the movie. By manipulating the displacement parameters, however, the display can be made either to appear as a square being transformed into a triangle or as a disappearing square and an appearing triangle. Of more practical import, the primacy of spatiotemporal continuity accounts for the fact that one's perception of an object can evolve over time. For example, a vehicle viewed from one's rearview mirror may appear as a police car when distant and as civilian automobile with a roof rack when near. While the identity ascribed to the object changes, its continuity as a single object remains stable. Indeed, an important aspect of the file metaphor is that information can be added as it

becomes available during the course of a perceptual episode. Importantly, because the information is indexed by location, preview benefits arising from the object token representational system should be location-dependent.

To determine whether the preview benefit depends on the continuity of object location, Pollatsek, Rayner, and Henderson (1990) used a modified version of the extrafoveal preview paradigm. The primary difference was that the new version had two objects in the preview display, side-by-side in the periphery. When the participants initiated a saccade, one of the preview objects was replaced with a target object and the other was replaced with a checkerboard mask so that only one nameable object remained. The location-dependency of the preview benefit was examined by manipulating the target object such that it occurred in the same or in a switched position relative to its position in the preview display. The greatest portion of the benefit observed in this version of the task was location-independent. That is, there was an advantage associated with having the target object in the preview display, but the additional benefit of having it remain in the same location was small. As a result, they suggested that the identification of objects from one fixation to the next was primarily facilitated by the activation of object representations stored in long-term memory.

While the Pollatsek et al. (1990) study implicated representations that are not referenced by location, Kahneman et al. (1992), found evidence for the primary involvement of spatially-indexed representations in the integration of information across disruptions of another kind. Their study was aimed not at the integration of information across eye movements but at the maintenance of an object's identity through change and motion. The logic and technique, however, were nearly identical. Whereas the Pollatsek

et al. (1990) experiments provided an index of the benefit of having the preview in the same versus a different location, Kahneman et al. (1992) established a benefit that is tied to having identity information associated with the same versus a different object. In their experiments, objects were defined as entities independent of their identities. This was accomplished by having square frames appear in one display with a letter in each. A linking display containing empty frames appeared in such a way as to produce the perception that the frames moved from one location to another. When the frames arrived at their final location, one of the preview letters was displayed in either the same or a different frame. Compared to a control condition, the letter was named faster if it appeared in the same object frame. There was little or no benefit derived from the mere presence of the letter in the other object frame, supporting the idea that the maintenance of object identity is accomplished through the reviewing of object files. Specifically, Kahneman et al. proposed that an object file is created during the initial view of an object and information is collected within the file as it becomes available, including a visual description of the object as it develops and the identity of the object once recognized. During subsequent views, the object file can be retrieved on the basis of its spatial and temporal position and target identification can then be speeded by reviewing the contents of the file.

The fact that performance in these two similar tasks favored two different representational systems warranted further investigation. It was possible that the discrepancy between the two studies could be accounted for by the fact that viewing in one was transsaccadic and the other was within-fixation; however, there were a number of methodological differences that might also have contributed to the differences in

results. To address these issues, Henderson and Anes (1994) put together a study that captured elements of the two previous approaches. Like the Pollatsek et al. (1990) study, they measured transsaccadic effects. However, like the Kahneman et al. (1992) study, they used letters in frames and a smaller stimulus set. In addition, the items were aligned vertically in the preview and target displays and the mask was eliminated. Kahneman et al. argued that the appearance of the mask in the switch condition of the Pollatsek et al. study might have generated the perception of motion. If this were the case, the observed effects would have to be considered object-specific. Finally, Henderson and Anes manipulated the number of task-relevant items in the preview display. While the target display always had a single to-be-named letter flanked by a plus sign, the preview display could have either two letters (a target letter and a flanking letter) or a letter and a plus sign. This manipulation was expected to provide converging evidence for the involvement of object files under the assumption that only the construction or reviewing of object files would be capacity-limited. Thus, if there was a location-dependent component to the effect, only it would be reduced by the additional item in the preview display. With these modifications in place, Henderson and Anes found both location-dependent and location-independent preview benefits: targets were named more quickly in the same versus switch conditions as well as in the switch versus control conditions. In addition, only the location-dependent benefit was reduced by having a task-relevant flanker object in the preview. When the preview was two letters as opposed to a letter and a plus sign, the object-specific benefit was reduced but the nonspecific benefit remained unchanged.

The Two Representational Systems Theory

On the basis of the findings discussed above, Henderson (1994) proposed a two-representational-systems theory to explain how information from one fixation might facilitate the identification of an object viewed in a subsequent fixation. On this account, two sources contribute independently to the integration process. First, the initial view of an object generates activity at a number of levels of representation in the system that supports object identification. As a result, the identification of target objects can be facilitated by the priming of visual descriptions stored in long-term memory, basic-level semantic categories, and/or the object's name. A second source of facilitation is derived by the construction and review of object files (Kahneman et al., 1992). However, because only the location-dependent component of the effect was modulated by the task relevance of the flanking object, either the construction or the reviewing of object files is resource-limited.

The generality of the two-representational-systems framework has subsequently been tested using pictorial stimuli. The pattern of results found using letters in frames was replicated by Henderson & Siefert (2001) using line drawings of objects: both location-dependent and location-independent benefits were found and only the location-dependent benefit was reduced by the presence of a meaningful flanker in the preview display. More telling, however, was the pattern of results observed in a second experiment that included a mirror reversal condition, a condition that manipulates visual but not semantic or name content. Here, the location-dependent benefit was reduced when the preview and target were mirror images, but the location-independent benefit was undiminished by this manipulation. This has been taken as rather strong evidence in

favor of two independently contributing representational systems. While the type representation would be abstracted to the identity or concept level, only the token representation would be expected to preserve detailed information about the form of the specific object.

The two-representational-systems theory did, however, change in terms of its alliance with the object file theory. Henderson and Siefert (2001) opted to use the term, token, to refer to the episodic representation of their theory because the pattern of results observed in the transsaccadic studies was not entirely consistent with the object file theory. In particular, the object file theory suggests that the object file is the representation that gets matched to the long-term representations during identification, and as a result, the priming of object types is mediated by the object file. In the two-representational-systems theory, however, a resource limitation is associated with the episodic representation but not with the priming of object types. To accommodate the transsaccadic data, the object file theory would have to be modified to include a limitation on the review of object files that is not imposed on their construction.

An additional finding that may be difficult to reconcile with the object file theory is the fact that the pattern of results found in the transsaccadic studies holds when the retinal events are simulated within a steady fixation (Henderson, 1994; Henderson & Anes, 1994). Because the objects are moved from the periphery to the center of vision while fixation is maintained, the spatial coordinates of the objects change whether the target object is in the same or a switched location. This finding suggests that a configural spatial code plays a role in the integration process, and it is unclear how such a coding scheme would play out in the object file theory.

Overview of Current Research

While the pattern of results found in the location-dependency experiments suggests that integration can occur within a system that codes object types as well as within a system that codes object tokens, integration can be taking place at a number of levels of representation within each of these systems. Indeed, the study by Pollatsek et al. (1984) suggested that visual information and the object's name are each maintained and integrated across eye movements. According to the two-representational-systems theory advanced by Henderson (1994; Henderson & Siefert, 2001), the presaccadic allocation of attention towards the objects in the periphery can result in activation at the level of stored visual descriptions, semantic categories, and/or object names. Residual activity in the object recognition system can then produce location-independent priming at each of these levels when the target is processed after the saccade. In addition, a small number of object tokens will be constructed before the eye movement. When spatial continuity is maintained across the saccade, the token is retrieved and the properties of the object are reactivated within the object recognition system, leading to a location-dependent source of facilitation. In sum, the research and theory to date suggest that the integration of information about real-world objects can range from fairly detailed visual descriptions of objects to their conceptual identities and names. The overarching goal of the present study was to determine the relative contribution of the varied levels of representation using full-color pictures of real-world objects in the extrafoveal preview paradigm.

One objective was to determine the kind of visual information that plays a role in transsaccadic object identification. While studies in the context of reading suggest that only abstract letter codes are integrated (e.g., Rayner et al., 1980), previous research

using pictures of objects suggests that visual features contribute to the integration process. Evidence for a visual component is given by an advantage for identical preview and target objects over that derived by objects that visually differ, such as mirror reversals and token substitutions (e.g., Henderson & Siefert, 2001; Pollatsek et al., 1984). The additional benefit for the identical preview is diagnostic because the two kinds of previews differ only in terms of their visual properties. Preview benefits could arise from higher levels of representation in both of these cases because the preview and target objects always share a conceptual identity and name; however, the advantage for the identical preview is thought to reflect a visual component because the contribution of the higher levels of representation is assumed to be equated. In an effort to systematically investigate the kinds of visual properties that are pertinent, Experiments 1 and 2 manipulated the visual similarity between preview and target objects. Of particular interest was the nature of the difference that was required to produce an identical preview advantage.

A second objective was to examine the effects of extrafoveal previews using a non-repeating stimulus set. To date, the studies that have investigated the transsaccadic integration of pictorial information have all employed relatively small sets of items that repeat within a session (e.g., Gajewski & Henderson, 2005; Henderson, 1992a; Henderson, Pollatsek, & Rayner, 1987; Pollatsek et al., 1984, 1990). That is, naming latencies for a given target object were measured for each participant as well as for all preview conditions employed. Initial studies examining transsaccadic integration in reading were met with criticism due to the fact that items came from a limited stimulus set (McClelland & O'Regan, 1981; Paap & Newsom, 1981). The argument was that

preview benefits could be driven by expectations derived from the repetition of stimulus items within an experimental session. Preview benefits were found to survive when repetition was controlled in the context of reading (Balota & Rayner, 1983), but this issue has not yet been addressed using pictorial stimuli.

Testing the generality of the preview benefits with pictorial stimuli in the non-repeated context is important for at least two reasons. First, repetition effects on preview benefits might be stronger with objects compared to words because object shapes are likely to be visible in the periphery due to the usefulness of lower spatial frequency information. As a result, familiarity with the objects could allow participants to bypass the normal identification process by simply associating a constellation of (context-specific) visual features with an object's name. If facilitation is limited to contexts where familiarity with the objects is high, it would constrain what could be said about the continuation of identification processing across saccades.

Second, because identification of the preview object itself has been shown to occur quite readily in the repeated-item context (Pollatsek et al. 1984), investigations using item repetitions may not be best suited for testing integration at visual levels of representation. As mentioned above, evidence for a visual component to transsaccadic integration is given by the identical preview advantage. While testing the generality of the preview benefits is interesting in its own right, examining the impact of visual differences in the non-repeated context should be most informative. That is, assuming the identical preview advantage reflects the occurrence of integration at a level of representation prior to identification, the paradigm should be most sensitive to these

differences when the probability of reaching identification before the saccade is minimized, as would be the case when objects are not repeated.

A third objective was to determine the relative contribution of visual versus identity and name levels of representation to the generation of preview benefits in the non-repeated context. Preview benefits observed in the extrafoveal preview paradigm suggest that information acquired from the periphery during one fixation can facilitate the identification of objects during a subsequent fixation; however, this facilitative effect could be driven almost entirely by the attainment of identification before the saccade. Experiment 3 addressed this issue in two regression analyses. In the first analysis, preview benefits were examined as a function of accuracy in an extrafoveal identification task, which required participants to identify objects based solely on a brief peripheral glimpse. If preview benefits are driven primarily by identification of the preview object before the saccade, extrafoveal identification should strongly predict the magnitude of the preview benefits observed, and there should be little or no benefit for objects that are rarely or never identified in the periphery. On the other hand, if preview benefits do not depend on identification, there should be facilitation even for objects that are rarely or never identified peripherally.

In the second analysis, the contributions of visual and higher levels of representation were quantified by a multiple regression model that simultaneously accounted for preview identification and visual similarity. Of particular interest was whether the effect of visual similarity on preview benefits would vary as a function of extrafoveal identification. One hypothesis was that the visual contribution to the effect should be greatest for objects that are not readily identified from the periphery. When

objects are readily identified from the periphery, the identity of the object should frequently be activated or retrievable upon completion of the saccade, potentially obscuring integration at visual levels of representation. If this is the case, the benefit of visual similarity should be greatest for objects that are at the low end of the extrafoveal identification scale. On the other hand, objects that are more readily identified from the periphery may better activate the store descriptions that support identification. If this is the case, the benefit of visual similarity should be greatest for objects that are at the high end of the extrafoveal identification scale.

EXPERIMENT 1

Experiment 1 employed the extrafoveal preview paradigm of Pollatsek et al. (1984), but with photo-realistic pictures of objects instead of line drawings. Each trial consisted of three displays, as depicted in Figure 1. First, a fixation display was presented, consisting of a fixation cross on the left-hand side of the screen and a square frame on the right-hand side of the screen. The participant began each trial with their gaze directed at the fixation cross. Second, a preview display was presented. The preview display was exactly the same as the fixation display except that an object (meaningful or meaningless) appeared within the frame on the right. Participants were instructed to shift their gaze toward the object in the frame as quickly as possible once it appeared. Third, a target display was presented once the eyes crossed a software-defined boundary. Target displays were configured the same as the preview displays, except that only meaningful objects would appear within the frame on the right. Participants named the target object as quickly as possible and their vocal response terminated the target display.

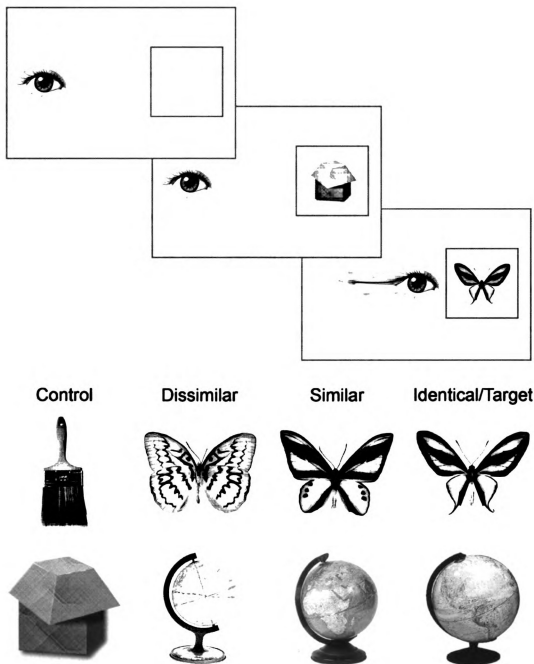


Figure 1. Schematic illustration of the displays presented in Experiment 1 (Top). During the first display, participants fixated a small plus sign on the left-hand side of the screen. A preview object was then presented in the second display and participants initiated a saccade toward the object. While the eyes were moving, the display was changed to present the target object. Participants named the target object as quickly as possible after the saccade. Example items for the meaningless-object, different-object, and identical preview conditions are also shown (Bottom). Full-color images were used in the actual experiment. The trial illustration is not shown to scale.

Experiment 1 was designed to satisfy three goals. The first goal was to test the generality of the preview benefit by using a non-repeating stimulus set. Each object was presented only once to each participant. Thus, there was no opportunity for participants to generate expectations concerning the stimulus set or to learn to associate particular visual features in the periphery with specific target objects.

The second goal was to examine the kind of visual information that is integrated across eye movements. Pollatsek et al. (1984) showed that a portion of the benefit is derived from the activation of visual features; however, the amount of visual feature overlap that is needed to produce facilitation remains an open question. In addition, the studies manipulating location suggest that two kinds of visual representations contribute independently to the integration process: spatiotemporally addressed episodic representations that are thought to include visual details associated with particular instantiations, and stored object descriptions that are considered more abstract. To examine the contribution of detailed versus abstract forms of visual representation, the visual similarity of preview and target objects was manipulated in four preview conditions (identical, visually-similar, visually-dissimilar, and control). Preview and target objects in the experimental conditions could differ visually but were always from the same basic-level category, and the objects were selected so as to be from the same approximate viewpoint (see Figure 1). As a result, they differed primarily in terms of surface-level features, such as color, texture, and the shapes of the parts. If preview benefits are driven by representations that preserve these properties, there should be a reduction in the magnitude of the preview benefit that corresponds to the reduction in the visual similarity between the preview and target objects. On the other hand, if preview

benefits are driven primarily by representations that are abstracted over these properties, the manipulation of visual similarity should have little or no effect.

The final goal was to determine the extent to which name priming contributes to performance in the non-repeated context. Pollatsek et al. (1984) found an inhibitory effect when the preview and target objects were from a different basic-level category: target objects were named more quickly in a control condition without a preview object than in the different-object preview condition. The inhibitory component was determined to reflect the availability of the preview object's name. When the preview object was closer to the point of initial fixation, the preview was more readily identified and name inhibition was elevated. In the present study, identification of the preview objects was assumed to be less frequent because each object appeared only once per session. Nevertheless, the contribution of name priming to the preview benefit is to some degree indicated by the amount of interference generated by the different object preview. Experiment 1 was divided into two subexperiments. In Experiment 1A, the control object for each target was from a different basic-level category than the target. In Experiment 1B, a meaningless non-object was used as a control. If name activation is a dominant component of the preview benefit in the non-repeated context, naming should be faster in the non-object control condition than in the different-object control condition, because only meaningful preview objects are expected to be associated with a name that would interfere with the naming of the target object.

Method

Participants. Thirty-two Michigan State University undergraduate students participated in the experiment for course credit, 16 each in Experiments 1A and 1B. All

participants had normal or corrected-to-normal vision and were naive with respect to the hypotheses under investigation.

Stimuli. The stimuli consisted of full-color pictures of real-world objects and a meaningless object. The meaningless object was created by taking a mottled color pattern and shading it to give it dimension. The shading was based on overlapping simple geometrical figures. Three exemplars of 60 object types were selected from the Hemera Photo-objects 50,000 Premium Image Collection. Objects were selected so that all the tokens within a category were from the same approximate point of view. The selection of real-world objects was based on two norming studies: the first rated pairs of object tokens for visual similarity, and the second rated the target objects for naming consistency (see Appendix A). Target, similar, and dissimilar objects were selected on the basis of the mean similarity ratings for each pair of object tokens so as to minimize the visual difference between the target and similar objects, and to maximize the visual difference between the target and dissimilar objects. The mean visual similarity score was reliably greater in the visually-similar condition (3.39) compared to the visually-dissimilar condition (2.19), $t(59) = 19.49, p < .001$, and all target objects were given the same name at least 75% of the time. An additional 15 objects were selected for the different-object preview condition. Examples for two object types are shown in Figure 1, and the visual similarity and naming consistency scores for the objects employed are presented in Appendix B.

Preview and target displays comprised an object centered within a square frame on the right-hand side of the screen and a fixation cross on the left-hand side of the screen. A total of 196 displays were generated using the three exemplars of 60 objects,

the 15 objects in the different-object preview condition, and the meaningless object. The object pictures were 5.5° in height and 5.1° in width on average at a viewing distance of 58 cm. The meaningless object was 5.8° in height and 6.4° in width. The frame subtended 8.8° vertically and horizontally. The fixation marker was vertically centered and 4.0° from the left-hand side of the screen. There was a 23.8° separation between the fixation marker and the center of the object frame on the right-hand side of the screen.

Apparatus. Stimuli were displayed at a resolution of 800 by 600 pixels by 24-bit color on a 19-inch Dell P991 monitor driven by a NVIDIA GeForce3 video graphics card with a screen refresh rate of 100 Hz. The room was illuminated by fluorescent overhead lighting. Eye movements were monitored using an ISCAN ETL-400 pupil and corneal reflection tracking system sampling at 240 Hz. The position of the right eye was tracked, though viewing was binocular. The eyetracker is accurate to within 0.25° of visual angle both horizontally and vertically. The computer changed the display contingent on detecting an eye movement that crossed an invisible boundary positioned 3.3° to the right of the fixation marker and 20.6° to the left of the center of the target objects. Display changes required a maximum of 20 ms and were accomplished during the saccade when vision was suppressed.

Stimulus presentation and response collection were controlled by E-Prime experimental software. Naming latencies were collected with a voice key provided by E-Prime. The eyetracker and display monitor were interfaced with a 2GHz, Pentium 4, microcomputer. The computer controlled the experiment and maintained a complete record of the position and time values for the point of regard, as well as time values for voice key events over the course of each trial.

Procedure. Upon arriving for the experimental session, each participant was seated comfortably. A forehead rest minimized head movements and maintained viewing distance. The session began with a generic object naming task to provide the experimenter an opportunity to adjust the sensitivity of the microphone. None of these objects were used in the actual experiment.

The eyetracker was calibrated at the beginning of the session and then checked between trials using the fixation display. Participants were asked to direct their eyes to the fixation marker and to the center of the object frame. If the calibration was satisfactory (plus or minus 0.5° from each of the positions), the participant was asked to direct their gaze toward the fixation marker to indicate readiness to begin. The experimenter then initiated each trial by pressing a silent button. The fixation display was replaced by the preview display and the participant immediately initiated a rightward saccade to the object centered within the frame. During the saccade, the preview display was replaced by the target display. The target display remained in view until the participant responded by naming the object as quickly as possible.

In both Experiments 1A and 1B, each participant named 60 objects. Trials were produced by the within-participant combination of 4 preview conditions: identical, visually-similar, visually-dissimilar, and control. For Experiment 1A, the controls were objects from a different basic-level category. For Experiment 1B, the control was always the non-object. Within each subexperiment, items were assigned to preview conditions via Latin square design so that each object appeared in each condition an equal number of times across participants. The order of object presentation (and hence the order of condition presentation) was determined randomly for each participant within each

subexperiment. Participants were assigned to subexperiment using a pseudorandom procedure; each participant took part in only one experiment. The entire session lasted approximately 30 minutes.

Results and Discussion

The mean naming latencies for this analysis appear in Table 1. Naming latencies were defined as the elapsed time between the crossing of the display-changing boundary to the onset of the vocal response. These means exclude trials in which the target object was named incorrectly, an anticipatory eye movement occurred (saccade latencies of less than 100 ms), and trials on which the naming latency was less than 200 ms or more than 3 standard deviations greater than the mean naming latency for that subject. Eliminated trials accounted for 10% of the data in Experiment 1A and 11% of the data in Experiment 1B. Saccade latencies were marginally slower in Experiment 1A (mean = 323 ms) than in Experiment 1B (mean = 269 ms), $F(1,30) = 3.00$, $MSE = 31,811$, $p = .09$, but did not differ across conditions in either experiment, $F < 1$, and, $F(3,45) = 1.725$, $MSE = 11,235$, $p = .18$, respectively. The source of the between-experiment difference in saccade latencies is unknown; however, the analyses that follow do not indicate a differential impact on the measures of interest.

The first question addressed in Experiment 1 was whether item-familiarity through repetition is a prerequisite for the observation of extrafoveal preview benefits on the identification and naming of real-world objects. Analyses of variance (ANOVAs) were performed on each subexperiment by participants (F_1) and by items (F_2) with preview condition as the within-participants and within-items factors respectively. There was an effect of preview condition in Experiment 1A, $F_1(3,45) = 8.53$, $MSE = 7,995$, $p <$

	Identical	Visually Similar	Visually Dissimilar	Maximally Dissimilar	Different- object Control	Non- object Control
Exp 1A	796 (46)	802 (29)	812 (31)		933 (29)	
Exp 1B	774 (24)	806 (28)	810 (27)			908 (31)
M	785 (25)	804 (20)	811 (20)			
Exp 2	677 (20)		690 (22)	721 (17)		785 (19)

Table 1. Mean Naming Latencies and Standard Errors (in milliseconds) for Experiments 1 and 2 by Preview Condition.

.001, $F_2(3,177) = 14.04$, $MSE = 18,491$, $p < .001$, and Experiment 1B, $F_1(3,45) = 12.07$, $MSE = 4,437$, $p < .001$, $F_2(3,177) = 12.14$, $MSE = 19,328$, $p < .001$. In both subexperiments, naming latencies were faster in all three experimental conditions than in the control conditions (all $ps < .01$). Naming latencies were slower overall compared to those typically observed using repeated items, but the magnitude of the preview benefits in the present study were as great or greater than those observed in previous studies (e.g., Henderson, 1992a; Henderson et al., 1987; Pollatsek et al., 1984, 1990). For example, across several experiments, naming latencies for the control conditions in the Pollatsek et al. (1984) study ranged from 681-787ms (compared to 908 ms in the present study), and the identical versus control preview benefits in their experiments ranged from 85-135 ms (compared to 136 ms in the present study). Thus, the results of Experiments 1A and 1B showed robust preview benefits, despite the fact that each participant saw each object only once.

The second objective was to determine the role of visual similarity in the generation of preview benefits. In Experiment 1A, there were no differences between experimental conditions when the control conditions were eliminated from the analyses ($ps > .82$). In Experiment 1B, the effect of preview condition was marginal by items,

$F_2(2,118) = 2.88$, $MSE = 17,682$, $p = .06$, when the control conditions were eliminated, but it was not reliable by participants, $F_1(2,30) = 1.64$, $MSE = 3,663$, $p = .21$. To determine whether the effect of visual similarity could be examined using the full power of both subexperiments, a mixed ANOVA was performed on the entire experiment with version included as a between-participants factor in the analysis by participants and as a within-items factor in the analysis by items. There was no effect of version, $F_1(1,30) = .03$, $MSE = 34,038$, $p = .87$, $F_2(1,59) = 2.17$, $MSE = 11,807$, $p = .15$, and the effect of preview condition did not differ between subexperiments, $F_1(2,60) = .27$, $MSE = 5,408$, $p = .77$, $F_2(3,177) = 1.20$, $MSE = 18,966$, $p = .31$. As within each subexperiment, the effect of preview condition was eliminated when the control condition was removed from the analysis, $F_1(2,60) = 1.036$, $MSE = 5409$, $p > .35$, $F_2(2,118) = 1.19$, $MSE = 8,937$, $p = .31$. However, because previous research has shown an advantage for identical previews, planned comparisons were conducted between the identical preview condition and the similar and dissimilar preview conditions. The contrast between the identical and dissimilar conditions was of particular interest because the items were selected so as to minimize the difference between the identical and similar conditions. Consistent with previous research, there was some indication that detailed visual representations contributed to performance: naming latencies were marginally faster in the identical preview condition relative to the dissimilar preview condition, $t(31) = -1.75$, $p = .09$, $t_2(59) = -1.68$, $p = .10$. The better part of the preview benefit, however, appears to be driven by a more abstract level of representation. The advantage of the dissimilar over the control condition (110 ms) was four times greater than the advantage of the identical over the dissimilar condition (26 ms). The fact that the visual effect was tenuous here suggests

that the visual differences between preview conditions were too small, and/or that the stimulus properties manipulated were largely inconsequential to the integration process.

The final objective of this experiment was to examine the role of name activation in the generation of preview benefits in the context of a non-repeating stimulus set. This issue was addressed by comparing naming latencies in the two control conditions. If name activation were a significant component of the effect, there should be interference when there is a mismatch between the name of the preview and the name of the target. While the mean naming latency was numerically greater in the different-object control condition than it was in the non-object control condition, this difference was not reliable, $t_1(30) = 0.605, p = .549, t_2(59) = 1.03, p = .31$. The result of this comparison suggests that name activation did not play as significant a role in the present study as it did in the Pollatsek et al. (1984) study. Assuming that name inhibition depends on preview identification, there are at least two differences between studies that could account for this difference. First, while full-color pictures of objects were employed here (as opposed to line drawings), the objects in the present study were displayed at a greater distance from the initial fixation. Second, the use of a non-repeating stimulus set was intended to reduce familiarity with the items. Each of these differences would be expected to reduce the frequency of preview identification, leading to a decreased contribution of name activation.

EXPERIMENT 2

While the identical preview condition showed an advantage relative to the dissimilar condition, the difference in naming latencies between these two conditions was small and statistically marginal. There are at least two potential reasons for the

tenuousness of the visual effect. First, visual differences between the preview and target object may have a lesser impact in the context of a non-repeating stimulus set. That is, while preview benefits measured as the difference between the identical and control conditions generalize to the non-repeated context, item-familiarity may be required to observe evidence of visual integration. This would be surprising, however, given that the lack of repetition was assumed to reduce the probability that the objects would be fully identified before the saccade, an assumption supported by the absence of name interference. If the contributions of the identity and name levels of representation are indeed reduced, one might expect a greater role for visual information in the integration process. Nevertheless, the overall slowing of naming in the infinite set paradigm could decrease its sensitivity to more subtle effects.

A second possibility, however, is that the visual differences between the preview and target objects in Experiment 1 were simply too small. The purpose of Experiment 2 was to extend Experiment 1 by employing a preview condition that introduced greater visual differences between the preview and target objects. The task and conditions employed were identical except that the visually-similar condition was dropped in favor of a very, visually-dissimilar condition (which will be termed the maximally-dissimilar condition), and only the non-object control was employed. The objects in the maximally-dissimilar condition were selected so as to be visually different and from a different viewpoint, visually different and mirror-reversed, visually different and rotated, or substantially different in another way (e.g., an apple chewed to the core as a preview for an unblemished apple, see Figure 2). If the limited effect of visual similarity in Experiment 1 should be attributed to the kind of visual differences employed, there

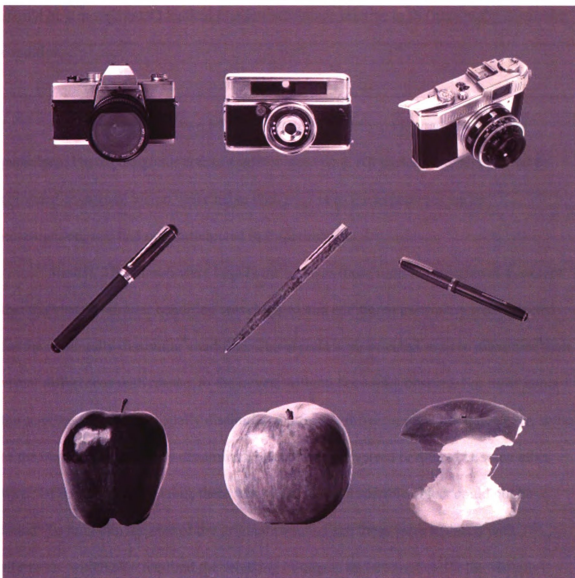


Figure 2. Example stimuli from Experiment 2. The columns from left to right correspond to the identical, visually-dissimilar, and maximally-dissimilar preview conditions. The objects in the identical and visually-dissimilar conditions were the same as those employed in Experiment 1. Most objects in the maximally-dissimilar condition were different exemplars from an absolutely different perspective (the camera), some were different exemplars mirror-reversed (the pen), and a few were substantially different in another way (the apple). Full-color images were used in the actual experiment.

should be a more robust identical preview advantage relative to the maximally dissimilar condition.

Method

Participants. Thirty-two Michigan State University undergraduate students participated in exchange for course credit or were paid. All participants had normal or corrected-to-normal vision, were naive with respect to the hypotheses under investigation, and had not participated in Experiment 1.

Stimuli. The stimuli were largely the same as those used in Experiment 1, except that the visually-similar condition was dropped and additional exemplars were selected for the maximally-dissimilar condition. The objects were selected so as to maximize their visual differences with respect to the targets without becoming obscure. For most items the selected object was visually dissimilar and from a different viewpoint; however, some of the objects were visually dissimilar objects mirror-reversed or altered in some other way. Of the original 60 items, there were 8 for which a suitable object could not be found. To maintain the size of the original stimulus set, these were replaced with alternates, which also required the selecting of targets and exemplars for the visually-dissimilar condition.

Apparatus and Procedure. The apparatus and procedure were identical to Experiment 1, except that a display containing the correct name of the object was presented on the computer screen after each object was named to facilitate on-line scoring by the experimenter, a set of 15 practice trials was administered immediately before each experiment began, and all participants were presented with the same non-object control condition. The entire session lasted approximately 30 minutes.

Results and Discussion

The mean naming latencies for this analysis appear in Table 1. As in Experiment 1, these means exclude trials in which the target object was named incorrectly, an anticipatory eye movement occurred (saccade latencies of less than 100 ms), and trials on which the naming latency was less than 200 ms or more than 3 standard deviations greater than the mean naming latency for that subject. Eliminated trials accounted for 16% of the data. Saccade latencies did not differ across the three experimental conditions (mean = 299 ms), $F(2,62) = 1.058$, $MSE = 1568$, $p = .35$, but were slower in the control condition (mean = 332 ms), $F(1,31) = 8.954$, $MSE = 2034$, $p < .01$.

Naming latencies were subjected to within-participant and within-item ANOVAs, which revealed reliable differences across the preview conditions, $F_1(3,93) = 18.20$, $MSE = 4,053$, $p < .001$, $F_2(3,177) = 13.00$, $MSE = 9,869$, $p < .001$. Naming latencies were faster in all three experimental conditions than in the control condition (all $ps < .01$), and there was an effect of preview condition when the control condition was removed from the analysis, $F_1(2,62) = 4.167$, $MSE = 3,953$, $p < .05$, $F_2(2,118) = 3.10$, $MSE = 7,684$, $p < .05$. Interestingly, although naming latencies in Experiment 2 were better than 100 ms faster here than in Experiment 1, the magnitude of the preview benefits were comparable: the differences between naming latencies in the identical and control conditions were 123 ms and 108 ms in Experiments 1 and 2 respectively. Of particular interest was whether there would be an identical preview advantage over one or both of the visually different preview conditions. Planned comparisons showed a 44 ms advantage for the identical preview over the maximally dissimilar condition, $F_1(1,31) = 9.198$, $MSE = 3410$, $p < .001$, $F_2(1,59) = 5.65$, $MSE = 8,135$, $p < .05$, but the 14 ms advantage over the visually

dissimilar condition was not reliable, $F_s < 1$. The fact that naming was faster in the identical preview condition than at least one of the visually different preview conditions dispels the idea that visual effects in the preview paradigm are strictly tied to repeated contexts. The clear failure to find an identical preview advantage over the visually-dissimilar condition, however, which was the same in both experiments, suggests that the marginal effect in Experiment 1 was spurious, and that the visual differences employed there were too subtle or the stimulus properties manipulated were inconsequential to the integration process.

Indeed, the overall pattern of results is indicative of the kind of visual information that is relevant to the integration process. Surface-level features, such as the object's color, texture and the shape of its parts, appear not to factor in prominently if at all: preview benefits in the visually-dissimilar preview condition were as great as those observed in the identical preview condition. It is important to note that the failure to find an effect of integration does not mean the preview and target objects in these conditions are indistinguishable across saccades. In a change detection task using the stimuli and display parameters of Experiment 1, differences between the visually-similar preview and the target could be detected 59% of the time, and differences between the visually-dissimilar preview and the target could be detected 87% of the time, both well above the 6% false alarm rate (for details, see Appendix C). Thus, it is not the case that the fleeting and poorly resolved extrafoveal retinal image renders the difference between these preview and target objects imperceptible.

While the failure to find an identical preview advantage over the visually-dissimilar condition suggests that the representations involved are abstracted away from

the surface-level properties, the advantage for the identical preview over the maximally-dissimilar preview suggests that the properties manipulated in that condition are important. Although the nature of the difference between these two conditions was somewhat varied and could be considered a matter of degree, the majority of objects represented a difference in viewpoint of one kind or another, whether by rotation, mirror-reversal, or by taking an entirely different perspective. The importance of maintaining an object's viewpoint across saccades should not be surprising given the empirical support for the view that object recognition is viewpoint-dependent (e.g., Tarr, Williams, Hayward, & Gauthier, 1998; see Tarr, 2003, for review). Transsaccadic integration cast in terms of object identification would be expected to reflect image properties that are captured by the object descriptions stored in long-term memory. If facilitation in the extrafoveal preview paradigm is in part determined by the priming of the object descriptions that support identification, it is the similarity or visual overlap between the preview object and the description that ultimately gets matched to the target object that should determine the magnitude of the benefit. The present findings could therefore be interpreted as consonant with image-based models of recognition (e.g., Poggio & Edelman, 1990; Riesenhuber & Poggio, 2000; Tarr & Bülthoff, 1998; Ullman, 1998), which suggest that the descriptions that support identification are two-dimensional images corresponding to a small number of familiar views of a given object.

The overall pattern of results is also indicative of the contribution of detailed versus abstract visual representations. Recall, the two-representational-systems theory (Henderson, 1994; Henderson & Siefert, 2001) suggests a contribution of object types and object tokens, with detailed visual information provided by the retrieval of tokens

and abstract visual information provided by the priming of types. Given that objects always appeared in the same location in this study (i.e., spatiotemporal continuity was maintained), preview benefits were expected to have a contribution from type priming as well as from token retrieval. In the present experiment, preview benefits were undiminished by the visual differences between preview and target objects in the visually dissimilar condition. If it were the case that these differences could not be discerned from the periphery, then the integration of visual information would simply reflect the precision of the episodic representation. Instead, these differences were readily noticed in the change detection task, which suggests that there is information captured by the episodic representation that has no bearing on the integration process.

EXPERIMENT 3

In contrast to the goals of Experiments 1 and 2, which was primarily to determine the kind of visual information that is relevant to the integration process, Experiment 3 was designed to determine the relative contribution of visual versus higher levels of representation, such as an object's conceptual identity and name.

In the present study, the use of a non-repeating stimulus set was assumed to reduce the probability that the objects would be fully identified before the saccade. The assumption that identification of the preview object factors in less prominently in the non-repeated context was supported by the lack of name interference. However, picture naming is generally held to be comprised of three relatively discrete stages: object identification, name activation, and response generation (Johnson, Paivio, & Clark, 1996). Thus, it is possible that attention to the preview object before the saccade led to identification but not name activation. Indeed, the results of Experiment 2 suggest a

substantial contribution of non-visual levels of representation. While there was a 44 ms identical preview advantage, the maximally-dissimilar preview condition generated a reliable 64 ms preview benefit measured relative to the non-object control, presumably driven by activation of the preview object's conceptual identity, and perhaps its name but to a lesser extent.

A question that arises, then, is whether partitioning the preview benefit into components that do and do not depend on the visual form of the preview object is equivalent to partitioning it into components that do and do not depend on preview identification. Theoretically, the identical preview advantage can be said to reflect only the contribution of visual information, because the preview objects in the identical and maximally-dissimilar conditions only differed visually. Similarly, the component that is unaffected by visual differences between preview conditions can be said to reflect the contribution of conceptual identity and name, because these were held constant across preview conditions. Assuming that only the contributions of conceptual identity and name depend on preview identification, attributing the identical preview advantage to the component of the preview benefit that does not depend on identification is perfectly reasonable. However, visual differences between the preview conditions may be accompanied by systematic differences in how readily the preview object can be identified. If the preview object tended to be more identifiable in the identical preview condition, the identical preview advantage would overestimate the visual component of the effect. An alternative approach is therefore needed to fully tease apart the visual and higher-level components, an approach that accounts for the identifiability of the preview object.

The approach employed in Experiment 3 was to determine the role of preview identification in the generation of preview benefits using a series of items-based regression analyses. The general idea was to examine preview benefits for objects in the extrafoveal preview paradigm as a function of their proportions correct in an extrafoveal identification task. The extrafoveal identification task was a modified version of the extrafoveal preview paradigm. Each trial began with a fixation display and was followed by a preview display that contained an object in a frame on the right-hand side of the screen. However, during the saccade the preview object was replaced with a question mark to cue participants to report the name of the object that appeared in the preview display. Thus, performance in this task was based solely on the information that could be acquired from a brief peripheral glimpse. The underlying assumption was that the proportion of participants who correctly identify a given object in the identification task reflects the probability that the object will be identified before the saccade in the preview paradigm.

In Experiment 3, there were two sets of analyses. The primary goal of the first analysis was to determine the predictive value of extrafoveal identification on the generation of preview benefits using a simple linear regression model. To accomplish this objective, naming latencies for 120 objects were collected in the identical preview and non-object control conditions of the extrafoveal preview paradigm, and identification accuracy was measured for each object in the extrafoveal identification task. Of particular interest were the slope and intercept terms given by the regression. Because the range of possible values on the extrafoveal identification scale extend from 0 to 1, the slope and intercept terms were considered indicative of the components that do and do not depend

on identification, respectively. That is, the magnitude of the preview benefit for objects that were never identified was given by the intercept with the Y axis, and the additional benefit for objects that were always identified was given by the slope.

An additional goal of the first analysis was to determine whether the extrafoveal identification task would predict the magnitude of the preview benefit beyond its relationship to foveal identification time. The extrafoveal identification measure reflects the probability that an object can be identified based on a brief peripheral glimpse, and objects that are more readily identified from the periphery would be expected to also be more quickly identified at fixation. Because the predictive value of extrafoveal identification could be completely tied to its relationship to foveal identification time, it is important to determine whether extrafoveal identification predicts the magnitude of preview benefits when foveal identification time is controlled for statistically. To accomplish this objective, naming latencies were collected for objects presented at the center of the screen with no eye movement needed, and the two predictors were examined in a hierarchical regression analysis with foveal naming time entered in the first step and extrafoveal identification accuracy entered in the second. If extrafoveal identification has unique predictive value, it should account for variance unaccounted for by the foveal identification time measure.

The second analysis also had two goals. The first goal was to determine whether viewpoint differences between the preview and target object are the primary visual determinants of the magnitude of the preview benefit using the regression approach. The objects in the maximally-dissimilar condition of Experiment 2 were primarily from different viewpoints, but this manipulation was not as systematic as would be ideal

because the viewpoint differences were created by rotation, mirror-reversal, as well as by taking an absolutely different perspective. The most powerful manipulation would be to cross viewpoint and surface-level differences in an experimental design, which would allow one to directly compare preview benefits for identical objects from different views with preview benefits for different objects from the same viewpoint; however, a sufficient number of object pairs from matching views could not be found in the database of photo-realistic objects employed in this study. The advantage of the regression approach is that the contribution of viewpoint and surface-level properties can be teased apart with fewer constraints on stimulus selection. To accomplish this objective, naming latencies were collected in a different-token version of the extrafoveal preview paradigm with the visual similarity of the preview and target items ranging from nearly identical and from the same viewpoint to appreciably different and from a different viewpoint. The object pairs were normed for object similarity, which was the visual similarity of the objects disregarding differences in viewpoint, as well as for viewpoint similarity, which was the similarity of the view of the objects disregarding differences in the objects themselves. If visual integration is based on representations that are abstracted over surface-level properties but not differences in viewpoint, the magnitude of the preview benefit should be predicted by the viewpoint similarity rating but not by the object similarity rating.

The second goal was to determine the contributions of visual and identity levels of representation to the integration process by examining the combined effects of visual similarity and extrafoveal identification in a multiple regression model. By including both predictors, the preview benefit can be partitioned into components that do and do not

depend on the visual form of the preview object as well as into components that do and do not depend on preview identification. In other words, because preview objects vary in terms of their similarity to the target object as well as their identifiability, the relative contribution of visual and higher levels of representation can be teased apart most effectively by accounting for each of these variables simultaneously.

An additional benefit of the multiple regression approach is that the effect of visual similarity can be examined as a function of extrafoveal identification by including a visual similarity x extrafoveal identification interaction term. Throughout the paper it has been assumed that the attainment of identification before the saccade would obscure the contribution of visual effects. In fact, part of the motivation for the use of a non-repeated stimulus set hinged on this idea. An alternate hypothesis, however, is that the objects that are more readily identified also better activate the stored visual descriptions. There is actually some indication that this might be the case. The Pollatsek et al. (1984) study included a retinal eccentricity manipulation, which had been shown to affect the identifiability of the preview object. Interestingly, there was a trend toward a greater identical preview advantage in the condition where the preview object had the higher probability of identification (see their Experiment 5). If this is the case here, the benefit of visual similarity should be greater for objects at the high end of the extrafoveal identification scale.

Method

Participants. One hundred and sixty-five Michigan State University undergraduate students participated in the experiment for course credit (25 in the extrafoveal identification task, 60 in the identical/non-object control version of the

extrafoveal preview paradigm, 30 in the different-token version of the extrafoveal preview paradigm, and 25 in each of 2 versions of the foveal naming task). All participants had normal or corrected-to-normal vision and were naive with respect to the hypotheses under investigation.

Stimuli. A set of 120 object pairs were selected for the study on the basis of a preliminary object-naming task. Target objects were selected so that the same name was generated by at least 85% of these participants. Each item was paired with another object from the same basic-level conceptual category. The pairs were selected by the experimenter with the goal of creating a stimulus set with a wide range of visual differences between the object pairs. The pairs of objects were then normed for visual similarity (see Appendix D), with ratings obtained on three scales: 1) object similarity, where participants were instructed to rate pairs based on the similarity of the objects themselves while disregarding differences in viewpoint, 2) viewpoint similarity, where participants were instructed to rate pairs based on the similarity of the viewpoint of the objects while disregarding differences in the appearance of the objects, and 3) image similarity, where participants were simply asked to indicate the visual similarity each of the two objects without further instruction. (The image similarity scale was included so that the relative importance of object- and viewpoint-similarity to judgments of visual similarity could be determined, but it was not a variable of primary interest.) The pairs of objects were divided into two sets prior to the collection of data. Objects in the first set were used as previews and targets in the identical condition and as previews in the different-token condition. Objects in the second set were used as targets in the different-token condition. The assignment of objects to sets was arbitrary but with a bias toward

putting the more canonical picture in the second set; however, none of the objects were obscure.

Procedure. Three tasks with slightly different procedures were employed in Experiment 3. The procedure for the extrafoveal preview paradigm was the same as in Experiment 2, except that number of trials doubled. The assignment of participants to conditions was based on two analyses of interest: one group of participants was presented with the identical and non-object control conditions and another group of participants was presented only with the different-token condition. Participants in the identical/non-object control group saw all 120 items with half of the items assigned to the identical condition and half assigned to the control condition. Objects were assigned to preview conditions via Latin square design so that each object appeared in each condition an equal number of times across participants. Participants in the different-token group saw all 120 preview and target pairs.

The procedure for the extrafoveal identification task was largely the same as that used for the extrafoveal preview paradigm, except participants were instructed to name the object that appeared in the preview display instead of the target display. During the saccade, the preview display was replaced by a display that contained a question mark centered within the frame. The question mark remained in view until the participant responded by naming the object, but speed of response was not emphasized. Separate groups of participants were used for each set of objects so that each exemplar appeared only once in a session.

The foveal naming task was comprised of 3 display events: A fixation display that was presented until the participant pressed the mouse button, an object display that was

presented until the onset of the voice response, and a scoring display that contained the correct name of the object. As with the extrafoveal identification task, separate groups of participants were used for each set of objects. The order of object presentation was determined randomly for all tasks and participants. The sessions lasted approximately 30 minutes.

Results and Discussion

Analysis 1: Preview Benefits by Extrafoveal Identification and Foveal Naming Speed

The first goal of Experiment 3 was to assess the relationship between preview benefits and extrafoveal identification using an items-based regression approach. To accomplish this goal, performance measures were acquired from three tasks: 1) naming latencies in the identical and non-object control conditions of the extrafoveal preview paradigm; 2) naming latencies in the foveal naming task; and 3) proportions correct in the extrafoveal identification task.

Naming latencies in the extrafoveal preview paradigm were defined as the elapsed time between the crossing of the display-changing boundary to the onset of the vocal response, and preview benefits were the differences in mean naming latencies between the identical and non-object control conditions by item. Consistent with Experiments 1 and 2, these means excluded trials in which the target object was named incorrectly, an anticipatory eye movement occurred (saccade latencies of less than 100 ms), and trials on which the naming latency was less than 200 ms or more than 3 standard deviations greater than the mean naming latency for that subject.

Naming latencies in the foveal naming task were defined as the elapsed time between the onset of the picture and the onset of the vocal response. As with the

extrafoveal preview paradigm, these means excluded trials in which the target object was named incorrectly as well as trials on which the naming latency was less than 200 ms or more than 3 standard deviations greater than the mean naming latency for that subject.

Performance in the extrafoveal identification task was measured as the proportions of participants that correctly identified each object on the basis of the preview alone. These proportions were based only on trials with saccade latencies of at least 100 ms. After computing means for the items, 3 were eliminated from the analysis because their mean naming latencies were more than 3 standard deviations greater than the mean of all the items. For the 117 remaining items, eliminated trials accounted for 13% of the data in the extrafoveal preview paradigm, 11% of the data in the foveal naming task, and 5% of the data from the extrafoveal identification task.

Figure 3 shows a frequency distribution for the extrafoveal identification task. The X axis represents the proportions of participants that correctly identified a given object, and the Y axis represents the number of objects that were correctly identified at a given rate. The mean proportion correct was 0.54. As can be seen in the figure, the entire range of identification scores was represented in the data: accuracy ranged from 0 to 100 percent.

It is important to note that the saccade latencies in the extrafoveal identification task were comparable to those observed in the extrafoveal preview paradigm. Because the preview object disappeared saccade-contingently in the identification task, participants might have been led to a strategy of delaying their saccades so as to allow attention to covertly dwell on the object. While saccade latencies were somewhat slower

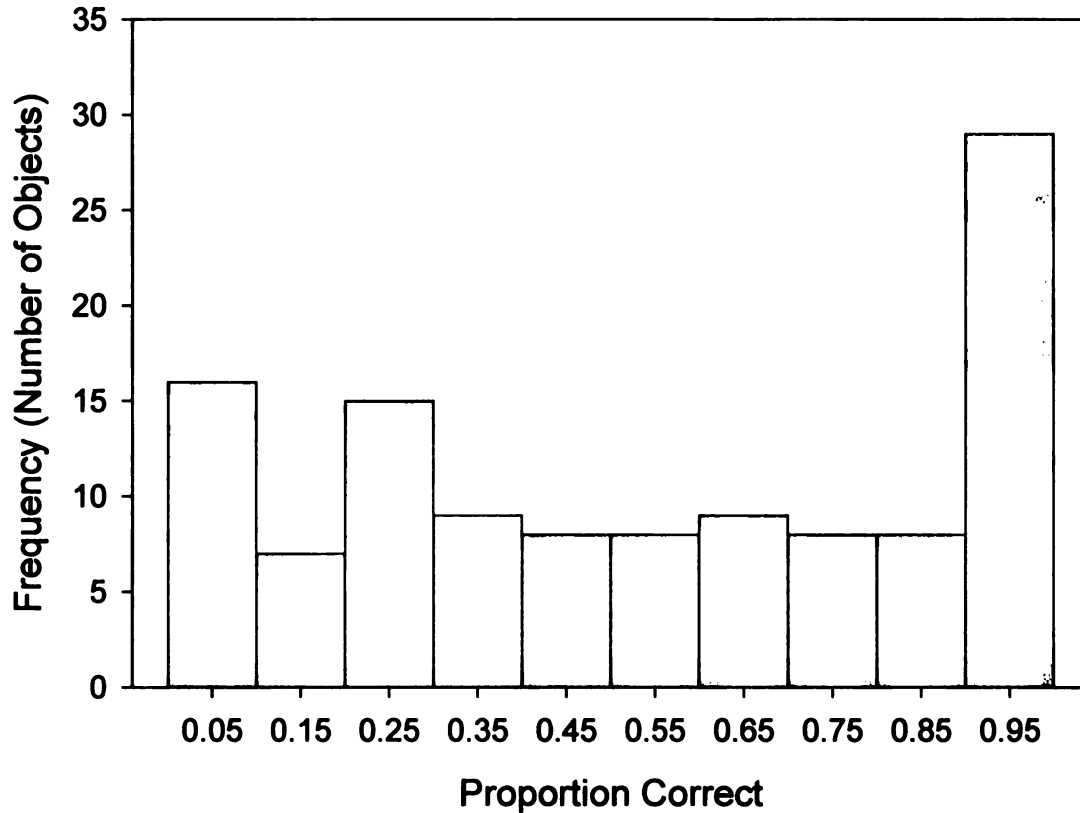


Figure 3. Frequency distribution for the proportions of participants who correctly identified a given object from the periphery in the extrafoveal identification task.

in the identification task (mean = 294 ms) relative to the identical preview condition of the preview paradigm (mean = 279 ms), $F(1,116) = 21.11$, $MSE = 601.07$, $p < .001$, the difference was small and latencies were slower in the non-object control condition (mean = 323 ms), $F(1,116) = 33.94$, $MSE = 1476.31$, $p < .001$. The similarity of the saccade latencies suggests that participants were not adopting different eye movement strategies in the two tasks. In contrast, participants were often surprised to learn that the timing of the display changes were under their behavioral control.

In the analysis of primary interest, preview benefits in the preview paradigm were examined as a function of the proportions correct in the identification task using a simple linear regression model. Figure 4 shows a scatter plot of the data with the regression line.

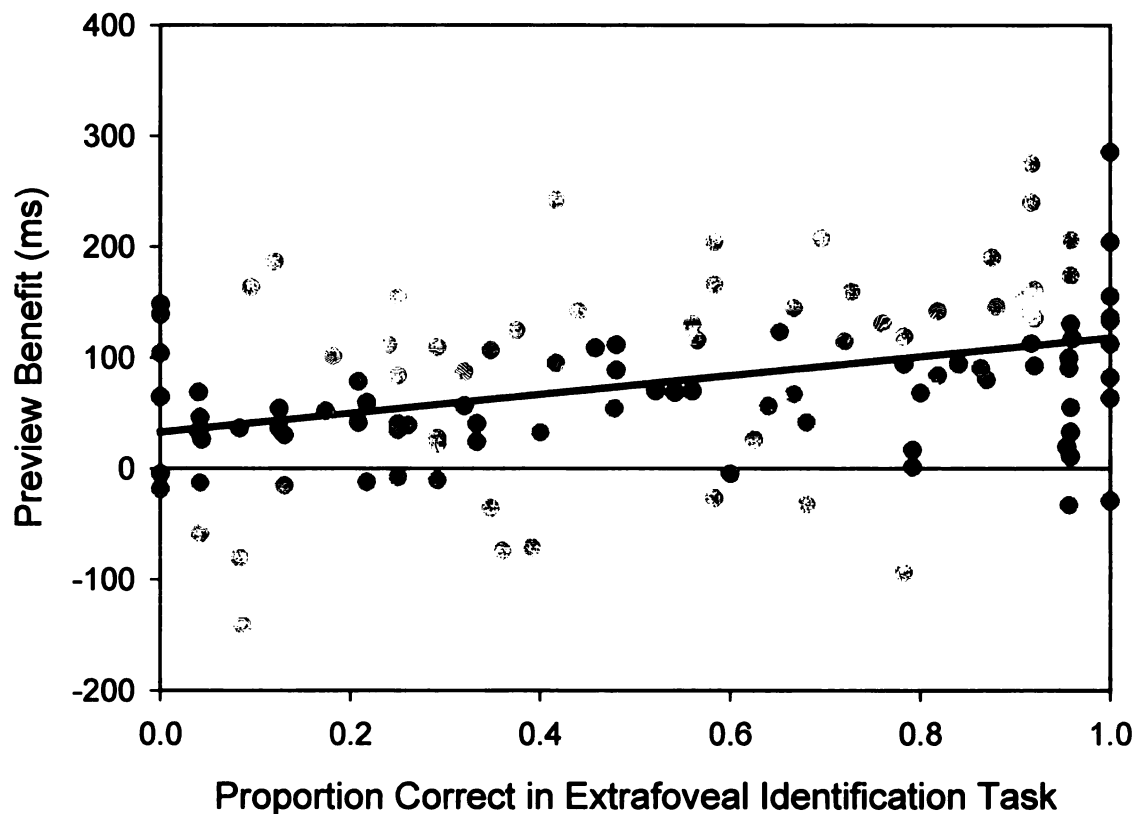


Figure 4. Preview benefits (in milliseconds) as a function of extrafoveal identification

The X axis represents the proportions correct in the identification task and the Y axis represents the corresponding means for the preview benefits. As can be seen in the figure, preview benefits rise with increasing identification accuracy; extrafoveal identification explained a significant amount of variance in preview benefits, $R^2 = .14$, $F(1,115) = 18.409$, $p < .001$. The fact that such a relationship is observed is not surprising; larger preview benefits would be expected for objects that are more readily identified from the periphery because more information would be available to contribute to target identification after the saccade. What is of interest is the degree to which preview benefits depend on extrafoveal identification. The most extreme possibility would be the case where facilitation occurs only when the preview object itself is identified. The

	1.	2.	3.	4.	5.
1. Extrafoveal Identification	--				
2. Foveal Naming	-.38*	--			
3. Naming Latencies (Non-object Control)	-.38*	.86*	--		
4. Naming Latencies (Identical Preview)	-.60*	.80*	.78*	--	
5. Preview Benefits	.37*	.03	.27*	-.40*	--

Table 2. Correlations for the Measures Employed in Experiment 3 (* $p < .01$)

present analysis does not favor such a conclusion, however, because the intercept term differed reliably from 0, $b_0 = 32.70$, $t(115) = 2.59$, $p < .05$, suggesting a 33 ms benefit for objects that are never identified from the periphery.

Although the data do not support complete dependence on extrafoveal identification, preview identification appears to play a prominent role in the generation of preview benefits. The slope term suggests an additional 85 ms benefit for objects that are always identified, $b_1 = 85.28$, $t(115) = 4.29$, $p < .001$. Given the equation that has emerged from the present analysis:

$$Y' = 85(\text{ExtraID}) + 33,$$

the component that does not depend on identification is a little more than a third the size of the component that does depend on identification. In contrast, the identical preview advantage of Experiment 2 was about two-thirds the size of the component that did not depend on visual form. The difference between experiments suggests that preview identification may have played a more prominent role in the generation of preview benefits than indicated by the identical preview advantage of Experiment 2.

The second goal of the present analysis was to determine whether the predictive value of extrafoveal identification extends beyond its relationship to foveal identification. Table 2 shows the correlations for the measures from the three tasks employed. As would

	<i>B</i>	<i>β</i>	<i>t</i>
Step 1			
Foveal Naming	.019	.029	0.31
Step 2			
Foveal Naming	.130	.197	2.14*
ExtraID	102.26	.446	4.84*

Table 3. Beta Weights From Regression Analysis I (* $p < .05$).

be expected, extrafoveal identification and naming latencies in the foveal naming task were reliably correlated, $r = -.38$, $p < .001$. It is not surprising that objects more readily identified from the periphery can be more quickly identified when foveated. The question is whether extrafoveal identification accounts for variance in preview benefits over and above the variance it shares with foveal naming speed. Naming latencies in the foveal naming task were highly correlated with those in the identical and non-object control conditions ($r_s \geq .80$, $p_s < .01$), which would be expected given that naming the objects at fixation is a component of performance in each of these conditions. What is interesting, however, is the fact that the correlation between the foveal naming and preview benefits was not reliable, $p = .76$. The failure to find an effect here suggests that preview benefits are independent of foveal naming speed, and as a result, there is no reason to expect the predictive value of the extrafoveal identification task to be mediated by foveal naming. Nevertheless, preview benefits were examined as a function of extrafoveal identification and foveal naming in a hierarchical regression analysis with foveal naming entered in the first step (see Table 3). When foveal naming and extrafoveal identification were both included in the model, both terms were reliable ($p_s < .05$). However, extrafoveal identification explained 17% of the variance in preview benefits unaccounted for by foveal naming, $\Delta R^2 = .17$, $F(1,114) = 23.43$, $p < .001$. Thus, extrafoveal identification has

unique predictive value; the probability that an object can be identified on the basis of a brief peripheral glimpse alone is not tied to the object's speed of foveal identification.

Analysis II: Preview Benefits by Extrafoveal Identification and Visual Similarity

The goal of this second analysis was to assess the integration of visual information across saccades by examining the effect of visual similarity on naming latencies in the preview paradigm. Of particular interest was whether preview benefits would depend more on viewpoint- than object-similarity, and whether the effect of visual similarity would depend on extrafoveal identification. To accomplish this goal, performance measures were acquired from three tasks: 1) naming latencies in the different-token condition of the extrafoveal preview paradigm; 2) naming latencies in the foveal naming task; and 3) proportions correct in the extrafoveal identification task. In the different-token condition, preview objects were taken from the first set of items and target objects were taken from the second. As a result, the second set of objects was employed in the foveal naming task and the first set was employed in extrafoveal identification task (as it was in the first analysis).

Preview benefits for this analysis were the differences in mean naming latencies between the different-token condition and the foveal naming task by item. As a result, naming latencies in the different-token condition were measured from the onset of the first fixation after the display-changing saccade. Consistent with all other analyses, these means excluded trials in which the target object was named incorrectly, an anticipatory eye movement occurred (saccade latencies of less than 100 ms), and trials on which the naming response (measured from the boundary crossing) was less than 200 ms or more

	1.	2.	3.	4.
1. Object-similarity	--			
2. Viewpoint-similarity	.17	--		
3. Image-similarity	.87*	.53*	--	
5. Preview Benefits	.10	.25*	.14	--

Table 4. Correlations of Preview Benefits with Visual Similarity Ratings (* $p < .01$).

than 3 standard deviations greater than the mean for that subject. To account for bad eye-tracking samples, trials were also eliminated if the difference between the saccade latency and the onset of the post-saccadic fixation was more than 2 standard deviations above or below the overall mean.

Naming latencies in the foveal naming task were defined as the elapsed time between the onset of the picture and the onset of the vocal response. As with the extrafoveal preview paradigm, these means excluded trials in which the target object was named incorrectly as well as trials on which the naming latency was less than 200 ms or more than 3 standard deviations greater than the mean naming latency for that subject. Saccade latencies averaged 289 ms in the different-token condition, which was not reliably different from the saccade latencies in the extrafoveal identification task (mean = 293 ms), $F(1,119) = 1.426$, $MSE = 626.17$, $p < .001$. The analysis was based on all 120 items. Eliminated trials accounted for 18% of the data in the extrafoveal preview paradigm, 10% of the data in the foveal naming task, and 5% of the data from the extrafoveal identification task.

The first goal of the analysis was to evaluate the three visual similarity scales by testing the correlation of each with preview benefits in the different-token condition (see Table 4). To begin, object-similarity and viewpoint-similarity were only marginally correlated ($r = .17$, $p = .07$), which shows that the objects differed independently on these

dimensions. Interestingly, image-similarity was more strongly correlated with object-similarity ($r = .87$) than with viewpoint-similarity ($r = .53$), $t(117) = 9.55, p < .01$, suggesting that differences between the objects themselves factor more strongly in the psychological assessment of visual similarity. Nevertheless, viewpoint-similarity was the only similarity scale that correlated with preview benefits ($r = .25, p < .01$). This finding provides converging evidence for the relative importance of maintaining viewpoint versus surface-level properties in the transsaccadic integration of visual information.

The second goal of the present analysis was to evaluate the combined roles of extrafoveal identification and visual similarity on the generation of preview benefits. Of particular interest was a) determining the relative contribution of visual and higher levels of representation by accounting for visual similarity and extrafoveal identification simultaneously, and b) determining whether the contribution of visual information changes over the levels of extrafoveal identification. As with the analysis above, the predictive values of the variables were evaluated using the hierarchical regression approach. Viewpoint similarity was the visual similarity scale employed in this analysis because it was the only scale that varied with the magnitude of the preview benefits. Extrafoveal identification was entered in the first step so that the additional contribution of viewpoint similarity could be evaluated in the second step (see Table 5). Extrafoveal identification again explained a significant amount of variance in preview benefits, $R^2 = .30, F(1,118) = 49.85, p < .001$, and the addition of viewpoint similarity in the second step improved the fit of the model, $\Delta R^2 = .04, F(1,117) = 6.46, p < .05$. However, the addition of the viewpoint-similarity x extrafoveal identification interaction term in the third step had no effect, $\Delta R^2 = .01, F(1,116) = 1.65, p = .20$, suggesting that the

	<i>B</i>	<i>β</i>	<i>t</i>
Step 1			
ExtraID	120.60	.545	7.06*
Step 2			
ExtraID	116.14	.525	6.92*
VP-similarity	14.75	.193	2.54*
Step 3			
ExtraID	33.24	.150	0.50
VP-similarity	3.27	.043	0.31
ExtraID x VP-similarity	22.13	.429	1.28

Table 5. Beta Weights From Regression Analysis II (* $p < .05$).

contribution of visual information was the same regardless of the identifiability of the preview object.

The results of the present analysis also provide a sharper characterization of the contributions of visual and higher levels of representation to the integration process. The model that best fits the data is given by the following equation:

$$Y' = 115(\text{ExtraID}) + 13(\text{VP-similarity}),$$

which was derived by including extrafoveal identification and viewpoint similarity as predictors. This model was run without a constant, however, because the constant term did not reliably differ from 0 in step 2 in the above regression, $b_0 = -8.01$, $t(117) = -.34$, $p = .73$. As a result, the coefficients differ slightly from those derived earlier. To quantify the contributions of preview identification and viewpoint similarity, consider the minimum and maximum values for each variable. The extrafoveal identification values ranged from 0 to 1, and the viewpoint similarity values ranged from 2.0 to 4.9. Plugging these values into the equation suggests a 0-115 ms component that depends on identification and a 26-64 ms component that depends on viewpoint similarity.

The present model is actually quite consistent with that generated in the first analysis. The difference between the maximum and minimum values for viewpoint similarity suggests a 38 ms advantage when preview and target objects are from the same viewpoint. Thus, the component that depended on visual properties of the object was one-third the size of the component that depended on identification. In the earlier analysis, the component that did not depend on identification was a little over one-third the size of the component that did depend on identification. Finally, the fact that there was no interaction is consistent with the idea that the visual and identity components contribute additively to the overall effect.

In sum, the results of Experiment 3 support the idea that the preview benefit can be partitioned into a component that depends on preview identification as well as a component that depends on the visual similarity of the preview to the target. While the better portion of the preview benefit reflects integration at the level of the object's conceptual identity, at least part of the effect is independent of identification, and the part that is independent of identification appears to depend on the maintenance of viewpoint across saccades but not the visual properties of the objects themselves.

General Discussion

The present study had three primary objectives. One objective was to determine whether preview benefits in the extrafoveal preview paradigm would generalize to the case where items are not repeated within a session. The facilitative effect's potential dependence on stimulus familiarity had already been ruled out in the context of reading (Balota & Rayner, 1983) but not in studies using pictorial stimuli. With pictures of objects, however, more useful information can likely be acquired from the periphery, and

as a result, the efficacy of repetition is likely elevated. Indeed, the fact that the identification of preview objects themselves occurs readily in the repeated-item context (Pollatsek et al., 1984) is consistent with the possibility that participants are able to bypass normal identification processes by associating context-specific visual features with an object's name. While the present study does not speak to the veracity of this hypothesis, it does indicate that preview benefits can be observed when context-specific associations are not given the opportunity to come into play. In Experiments 1 and 2, preview benefits for identical previews measured relative to the control conditions were all in excess of 100 ms. The fact that preview benefits were of comparable magnitude to those observed using repeated stimulus sets despite the fact that there were considerable between-group differences in the overall speed of naming indicates that the facilitative effect of the extrafoveal preview is quite robust. This finding is important for the continued-processing framework for transsaccadic integration, particularly the idea that transsaccadic integration can be thought of as object identification in the context of eye movements. If facilitative effects were only observed for familiar stimuli, it would suggest that integration is not a component process of normal identification, at least during one's initial encounter with an object.

A second objective was to determine the kind of visual information that is integrated across saccades, particularly in the context of a non-repeated stimulus set, where the contribution of the higher levels of representation was expected to be minimized. The benchmark for the integration of visual information was the advantage for the identical preview. If a preview object visually differs from the target object but produces the same facilitation as that produced by the identical preview, then the

properties that were not held constant cannot be said to play a role in the integration process. In contrast, if the difference between the preview and target diminishes the preview benefit, then the properties that were not held constant must be relevant. In Experiment 1, preview benefits were largely unaffected by visual differences between the preview and target objects, but the objects were always from the same approximate viewpoint. In Experiment 2, an identical preview advantage was observed relative only to the maximally-dissimilar condition. While the differences between previews and targets were somewhat varied in this condition, the majority reflected a change in viewpoint of one kind or another in addition to the more surface-level differences manipulated in the other dissimilar condition, such as color, texture and the shapes of the parts. The importance of viewpoint was reinforced in the second regression analysis of Experiment 3, which showed a relationship between preview benefits and viewpoint-similarity but not object-similarity. Thus, the results of this study indicate that surface-level properties of the objects play at best a minor role in the integration process. The visual properties that are integrated, however, are the properties that are common to the identical, visually-similar, and visually-dissimilar conditions, and differ between these and the maximally-dissimilar condition, such as the object's outline shape and/or its overall volumetric shape abstracted away from surface-level details.

An alternate way to frame the role for visual differences in the magnitude of the preview benefit is more quantitative than is suggested above. That is, perhaps preview benefits simply diminish when the visual differences between the preview and target are big. In the maximally-dissimilar condition, previews and targets differed in terms of viewpoint in addition to the surface-level differences that were manipulated in the other

conditions, and the viewpoint changes alone would be expected to alter the image greatly. However, at least two arguments can be made against the idea that the amount of change is the determining factor. To begin, the visual differences in the visually-dissimilar condition were readily noticed in a transsaccadic change detection task. If the magnitude of the preview benefit was in some way tied to the salience of the change, then the identical preview advantage would be about as robust in the visually-dissimilar condition as it was in the maximally-dissimilar condition. A stronger case is perhaps provided by the visual similarity analysis in Experiment 3. If preview benefits were affected only by the amount of change in the image, the image-similarity scale would be expected to correlate with preview benefits more strongly than either of the other two similarity scales. However, this outcome was not observed: viewpoint-similarity was the only similarity scale that correlated reliably with preview benefits. Ruling out the amount-of-change hypothesis is important because it would likely suggest that the visual effects are in some way artifactual. The visual source of facilitation has been cast as residual activation in the object recognition system (Henderson & Siefert, 2001). As a result, the kind of visual differences between preview and target that matter are expected to correspond to the kind of information is that captured by the representations that support identification. In contrast, it is unclear how the amount of change by itself would fit into an object recognition framework. Instead, it would best be explained as a general disruptive effect that occurs when a change is noticed.

A third objective was to determine the relative contribution of visual versus identity and name levels of representation to the generation of preview benefits. In Experiment 1, different-object and non-object control conditions were employed to assess

the contribution of name activation in the context of a non-repeated stimulus set. Previous research using a repeated stimulus set showed an effect of interference when the preview and target objects had different names (Pollatsek et al., 1984). This deficit suggested that the name of the preview object was activated and carried over into the processing of the target. In the present study, the naming of target objects was not statistically slower when the preview was a different object with a different name than when it was a meaningless non-object without an associated name. This finding suggests a reduced contribution of name activation when familiarity with the items is not developed through repetition, presumably due to a corresponding reduction in the frequency of preview identification.

While the lack of name interference indicated a reduced role for name activation, performance in the preview paradigm and the identification task suggested a contribution from the identity level of the representation. Although the rate of extrafoveal identification was not nearly as great as observed in the repeated context (Pollatsek et al., 1984), participants were able to identify the objects better than 50% of the time on average. In addition, a sizeable preview benefit was observed in Experiment 2 even when the visual differences between previews and targets were greatest. Thus, a primary issue in the present study was the extent to which preview benefits depend on the identification of the preview object itself.

In Experiment 3, the role for preview identification in the generation of preview benefits was examined in a series of items-based regression analyses. Preview benefits increased as a function of extrafoveal identification, but the facilitative effect was present for objects that were rarely or never identified on the basis of the peripheral glimpse alone. Thus, it cannot be said that preview benefits depend entirely on identification

before the saccade. The analyses do, however, suggest a prominent role for preview identification. While the non-zero intercept provided support for the idea that there was a component that does not depend on identification, this component was relatively small, amounting to about a 33 ms effect. Partitioning the preview benefit into components that do and do not depend on extrafoveal identification indicates a much larger (85 ms) component that does depend on identification. A similar decomposition was provided by accounting for visual similarity and extrafoveal identification simultaneously. It is important to note that partitioning the effect in this manner does not mean that identity priming is always the greater source of facilitation. If the preview object is not identified on a given trial, facilitation would be provided exclusively by the priming at the visual level of representation. Indeed, the fact that the viewpoint similarity x extrafoveal identification interaction did not approach statistical significance is consistent with the idea that priming at the identity level of representation contributes independently with priming at the visual level of representation.

In sum, previous research has suggested that there is a visual component to the integration process when pictures of objects are used as stimuli (Henderson & Siefert, 2001; Pollatsek et al., 1984; Pollatsek et al., 1990); the present study not only provides additional support for this conclusion but also indicates that integration is not dependent on the attainment of identification before the saccade. These findings support the idea that transsaccadic integration can be conceptualized as the continuation of processing across saccades, and suggest that integration is a component process of object identification in the context of eye movements. The term, transsaccadic integration, is more intimately tied to the classic notion that snapshots are merged together than it is

with the idea that the results of processing are combined across saccades. As a result, the integration problem has only loosely been tied to the problem of object recognition. From the continued-processing perspective, however, generating hypotheses about the kind of information integrated requires knowledge on the nature of the representations and processes involved in object identification. Since the problem of object recognition is not one that can be considered solved, the best that one can hope for is that theory and research from the two domains can be mutually informative. From this perspective one might ask: How does the present study contribute to an understanding of the problem of object recognition?

Aside from the suggestion that object identification can be thought of as a process that bridges discrete visual samples of the world, the most obvious contribution is derived by the nature of the visual mismatch between the preview and target that is required to observe an identical preview advantage. If the visual source of facilitation can indeed be cast as residual activation in the object recognition system (Henderson & Siefert, 2001), then the kind of visual differences that matter should correspond to the kind of information is that captured by the representations that support identification. In other words, the preview benefit should depend on how well the preview primes the representation that ultimately gets matched to the target after the saccade. As discussed above, preview benefits were robust to surface-level differences but not to differences in viewpoint. This finding is most consonant with the view-based approach, which suggests that the object recognition system encodes information about objects as viewed from particular vantage points (e.g., Poggio & Edelman, 1990; Riesenhuber & Poggio, 2000; Tarr & Bülthoff, 1998; Ullman, 1998). An object-centered approach, such as that

employed in Biederman's (1987) "Recognition-By-Components" model, which suggests that the object recognition system encodes objects in terms of their volumetric parts and the relations between the parts, could account for the viewpoint effect if the preview objects were systematically more difficult to decompose into parts when the viewpoint of preview and target differed maximally. This possibility seems remote, however, given that the objects employed in the present study were not depicted from obscure perspectives.

Another aspect of the current study that would seem particularly important for researchers studying object recognition is the fact that the objects could be so readily identified from the periphery. Although the lack of repetition decreased the frequency of extrafoveal identification overall, 29 objects were correctly identified 90% of the time or better. This occurred despite the fact that objects subtended around 5° of visual angle and were presented greater than 20° into the periphery. This finding would seem to place constraints on the kind of information that is needed to identify an object. Interestingly, the idea that identification can be based on visual input that is relatively coarse is reflected in a recent proposal suggesting that entry-level object recognition could be supported by low spatial scale images, something like blurry silhouettes of objects (Tarr, 2003). The idea is based largely on the possibility that fully-detailed images might not be best suited for perceptual categorization. In particular, the details associated with complete images are posited to reduce the similarity between exemplars of a category thereby making it difficult to map multiple exemplars to a particular class. The proposal is supported by the fact that recognition performance based on silhouettes can be as good or better than performance based on shaded renderings of objects (Hayward, 1998). In

addition, computational and behavioral studies indicate that silhouettes provide enough information to discriminate between classes of very similar objects, such as between dogs and cats (Cutzu & Tarr, 1997; Eimas & Quinn, 1994; Quinn, Eimas, & Tarr, 2001).

Present support for this idea comes from a subsidiary experiment using the extrafoveal preview paradigm and silhouettes generated for the objects of Experiment 1. Here, a reliable 43 ms preview benefit was observed relative to the non-object control condition (see Appendix F), suggesting that outline shape is an important property, a property that changes with changes in viewpoint. Importantly, it suggests that information about outline shape can be readily acquired from the periphery and used to support target identification. A question that might need to be asked is whether blurred extrafoveal retinal images can support the decomposition of objects into volumetric parts as suggested by Biederman's (1987) model.

Having considered the implications for the problem of object recognition, the discussion turns now to the present study's contribution to current theory of transsaccadic integration. Previous research has suggested that the integration of information about objects takes place at various levels of representation, ranging from detailed visual representations to the representation of the object's name (Henderson, 1994; Pollatsek et al., 1984). The overarching goal of this study was to sharpen theory not only by identifying the visual properties that are integral to the integration process but also by determining the relative contributions of the varied levels of representation.

To begin, the two-representational-systems theory proposed by Henderson (1994; Henderson & Siefert, 2001) suggests that the type and token systems contribute independently to the integration process. Prior to the execution of a saccade, attention is

directed toward the upcoming saccade target, and with that shift of attention, processing of the soon-to-be-fixated object begins. Facilitation of identification from within the type system is provided by activation at visual, identity and name levels of representation, depending on the level of processing achieved before the saccade landing. The present study indicates that objects differ widely in terms of how readily they can be identified from the periphery. These differences are likely tied to a number of factors, such as the object's divergence from the canonical view (Palmer, Rosch, & Chase, 1981) as well as the individual's personal history with objects of that kind. When the saccade target is fixated, identification is speeded by the combining of new and residual activation within the object type system. If the object cannot be identified from the periphery, priming will occur only at the level of the object's stored description. However, if the object can be identified from the periphery, priming will occur additionally at the level of the object's identity and perhaps its name. If spatiotemporal continuity is maintained across the saccade, facilitation of identification can also be provided by the token system. This source of facilitation is generated by the retrieval of the spatially-indexed episodic representation of the object created before the saccade. According to the theory (see Henderson & Siefert, 2001), retrieval of the token reactivates the properties of the object within the type system, and this reactivation combines with new activation in the exact same way that residual activation combines with new activation during the priming of types. That is, although the token is considered a more visually detailed representation, its influence on identification is mediated by the type system. The present data lend support to this idea. Preview benefits were largely undiminished by surface-level differences, suggesting a contribution from representations that are abstracted away from these

properties. However, these very same differences could be noticed in a change detection task, suggesting that the episodic representation provides details that are not relevant to the integration process.

Conclusion

To summarize, a number of conclusions can be reached on the basis of the present study. First, preview benefits have been shown to generalize to the case where items are not repeated. Second, integration is largely but not completely driven by identification of the object before the saccade. Third, the visual representations that are involved in the integration process are abstracted away from surface-level properties but not viewpoint. Fourth, the visual component of the preview benefit is independent of identification, suggesting that priming at the identity level of representation contributes additively with priming generated at the visual level of representation. Finally, the study validates the idea that the transsaccadic integration should be thought of as a component of object identification, and that the extrafoveal preview paradigm could be used as a tool to leverage theory in that domain.

APPENDIX A

Method for Visual Similarity Norm of Experiment 1

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

Stimuli. The stimuli consisted of full-color pictures of real-world objects selected from the Hemera Photo-objects 50,000 Premium Image Collection. The pictures were comprised of 140 object types: 4 exemplars of 135, and 3 exemplars of 5 for a total of 555 images. Objects were selected so that all the tokens within a category were from the same approximate point of view. A display was generated for all possible within-category object pairings: 6 for the 135 with 4 exemplars and 3 for the 5 with 3 exemplars for a total of 825 new images. The displays were each comprised of a neutral gray background, the trial list number, and two objects positioned side-by-side around the center of the display. The objects were of the same pixel dimensions as those employed in Experiment 1, and the displays were projected so as to subtend about the same number of degrees of visual angle at the average viewing distance.

Apparatus and Procedure. The visual similarity norm was conducted in a classroom by projecting the images on a screen via LCD projector. Each object pair object was presented in a random order (determined in advance of the session) for a duration of 5 seconds. A warning tone sounded one second before each display was terminated. All pairs of objects were rated by each participant on a 5-point scale with 5 indicating the highest degree of visual similarity and 1 indicating the lowest degree of

visual similarity. Responses were indicated by bubbling in scantron sheets, and the raw data was compiled by the optical scanning service offered by the Scoring Office at Michigan State University. The study was run in 11 sessions, and each session lasted approximately 90 minutes.

Method for Name Consistency Norm

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

Stimuli. The stimuli were 226 objects selected from the set employed in the visual similarity norm. The objects selected were candidate target objects for the extrafoveal preview paradigm (i.e., the objects that would ultimately be named). Because two similarity schemes were considered, one where the identical and similar previews were as similar as possible and one where the differences between similarity conditions were about equal, 86 out of the 140 object types required 2 exemplars in this study. The displays generated were each comprised of a neutral gray background, the trial list number, and one object at the center of the display. The objects were of the same pixel dimensions as those employed in Experiment 1, and the displays were projected so as to subtend about the same number of degrees of visual angle at the average viewing distance.

Apparatus and Procedure. The name consistency norm was conducted in a lecture hall by projecting the images on a screen via LCD projector. Each object pair object was presented in a random order (determined in advance of the session) for a duration of 6 seconds. A warning tone sounded one second before each display was terminated.

Participants were instructed to generate a name for each object and write it down on numbered score sheets that were provided by the experimenter. The study was run in 1 session that lasted approximately 35 minutes.

Results

Naming consistency was defined as the frequency of the most frequent response divided by total number of responses. Of the 140 object types employed in the visual similarity norm, 60 were selected based on a name consistency criterion of 0.75. Target, similar, and dissimilar objects were then selected on the basis of the mean similarity ratings for each pair of object tokens so as to minimize the visual difference between the target and similar objects, and to maximize the visual difference between the target and dissimilar objects. The mean visual similarity score for the sixty selected objects was reliably greater in the similar condition (3.39) compared to the dissimilar condition (2.19), $t(59) = 19.49$, $p < .001$. The object names, the visual similarity values for the similar and dissimilar previews, and the naming consistency rating for the targets are all presented in Appendix B.

APPENDIX B

Name	Similarity 1	Similarity 2	Name Consistency
Accordion	3.8	1.9	0.89
Apple	4.0	2.4	0.99
Ball	3.7	2.0	0.76
Basket	3.0	1.9	0.99
Bell	3.5	2.4	1.00
Binoculars	3.6	2.1	0.99
Boots	3.2	2.3	0.92
Bowl	3.4	1.6	0.89
Bullet	3.5	1.7	0.97
Butterfly	4.0	1.8	1.00
Cake	3.0	1.5	0.96
Calculator	3.1	2.5	1.00
Camera	3.0	2.4	0.99
Cane	3.6	2.5	0.90
Chair (1)	3.0	2.0	0.97
Chair (2)	3.5	2.3	1.00
Doll	2.8	2.0	0.92
Donut	3.9	2.0	0.97
Earphones	3.2	3.0	0.96
Earrings	2.4	1.8	0.93
Fan	3.7	2.2	0.86
Feather	3.2	1.9	1.00
Fire hydrant	3.7	1.9	0.91
Fireplace	2.9	2.0	0.92
Fish	3.7	2.3	0.87
Fork	3.4	2.1	1.00
Frog	3.3	1.8	0.89
Globe	3.3	2.3	0.97
Grapes	4.2	2.9	0.92
Guitar	4.0	2.6	0.94
Hammer	4.0	2.8	0.96
Hanger	3.6	2.7	0.92
Key	3.3	1.6	0.93
Lamp	3.9	2.1	0.90
Leaf	4.1	2.0	0.90
Light bulb	3.5	2.2	0.92
Lock	3.1	2.3	0.89
Medal	4.1	3.4	0.90
Microphone	3.2	2.2	0.79
Microscope	3.9	2.6	0.94
Mushroom	2.6	2.2	1.00
Pear	3.9	2.7	0.99
Pen	3.9	2.1	0.97

Pipe	3.9	3.0	0.99
Printer	3.2	2.6	0.94
Purse	2.9	1.4	0.96
Roller blade	3.0	2.7	0.93
Screw	3.1	2.0	0.90
Skateboard	3.5	2.6	1.00
Spoon	3.5	1.9	0.86
Stool	3.1	2.0	0.97
Sunglasses	3.0	1.9	0.94
Sword	3.6	1.6	0.92
Telescope	2.6	2.1	0.97
Tire	3.2	1.8	0.97
Typewriter	3.0	2.0	0.94
Vase	2.7	2.4	0.99
Walkie-talkie	3.7	2.5	0.90
Wheelchair	3.6	2.2	0.97
Wreath	2.4	1.8	0.82

APPENDIX C

Method for Change Detection Task

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

Stimuli. The stimuli employed were taken from the identical, visually similar, and visually dissimilar preview conditions of Experiment 1.

Apparatus and Procedure. The apparatus and procedure were the same as Experiment 1, except the naming response was replaced with a same-different judgment and the control conditions were not used. Participants began each trial with fixation on a cross on the left-hand side of the screen, and directed their gaze toward the object when it appeared in the peripheral frame. A saccade-contingent display change was then initiated so that upon completion of the saccade the object was the same, visually-similar, or the visually-dissimilar. Participants were instructed to indicate with a button press whether the preview and target objects were the same or different. Objects were assigned to conditions via Latin square design so that each object appeared in each condition an equal number of times across participants. The order of object presentation (and hence the order of condition presentation) was determined randomly for each participant. The entire session lasted approximately 20 minutes.

Results

Accuracy was higher in the visually-dissimilar condition (mean = .87) than in the visually-similar condition (mean = .59), $F(1,8) = 19.52$, $MSE = .02$, $p < .01$, but both were well above the false alarm rate (mean = .06), $ps < .001$

APPENDIX D

Method for the Visual Similarity Norm of Experiment 3

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

Stimuli. A set of 120 object pairs was selected from the Hemera Photo-objects 50,000 Premium Image Collection on the basis of a preliminary object-naming task. Target objects were selected so that the same name was generated by at least 85% of these participants. Each item was paired with another object from the same basic-level conceptual category. The pairs were selected by the experimenter with the goal of creating a stimulus set with visual differences ranging from nearly identical and from the same viewpoint to visually different and from a different viewpoint. The displays for the 120 pairings were each comprised of a neutral gray background, the trial list number, and two objects positioned side-by-side around the center of the display. The objects were of the same pixel dimensions as those employed in Experiment 3, and the displays were projected so as to subtend about the same number of degrees of visual angle at the average viewing distance.

Apparatus and Procedure. The visual similarity norm was conducted in a classroom by projecting the images on a screen via LCD projector. Each object pair object was presented in a random order (determined in advance of the session) for a duration of 10 seconds. A warning tone sounded one second before each display was terminated. All pairs of objects were rated for visual similarity by each participant on three 5-point scales with one being least similar and 5 being most similar. The three

scales were: 1) object similarity, where participants were instructed to rate pairs based on the similarity of the objects themselves while disregarding differences in viewpoint, 2) viewpoint similarity, where participants were instructed to rate pairs based on the similarity of the viewpoint of the objects while disregarding differences in the appearance of the objects, and 3) image similarity, where participants were simply asked to indicate the visual similarity each of the two objects without further instruction. (The image similarity scale was included so that the relative importance of object- and viewpoint-similarity to judgments of visual similarity could be determined, but it was not a variable of primary interest.) Responses were indicated by circling numbers on score sheets that were provided by the experimenter. The study was run in 13 sessions, and each session lasted approximately 30 minutes. The mean similarity ratings for all 120 object pairings is presented in Appendix E.

APPENDIX E

Object Name	Object Similarity	Viewpoint Similarity	Image Similarity
Apple	4.2	4.5	4.1
Backpack	3.4	3.4	3.2
Bagel	3.5	4.7	3.6
Balloons	3.3	4.4	3.3
Banana	3.3	2.3	2.9
Basket	3.7	4.8	3.9
Battery	3.5	4.8	3.5
Bear	3.1	2.0	2.9
Bed	2.8	2.7	2.8
Bell	3.6	4.8	3.6
Belt	2.9	3.5	3.1
Bench	2.8	2.4	2.6
Bib	3.0	4.7	3.1
Binoculars	3.8	4.7	3.8
Blender	3.2	3.1	3.1
Boat	4.1	2.8	3.9
Books	3.0	2.2	2.8
Bowl	3.5	4.9	3.7
Brush	2.8	4.8	3.0
Bus	3.1	4.3	3.4
Butterfly	3.9	4.8	4.1
Button	3.8	4.4	3.9
Cake	3.6	4.6	3.7
Calculator	2.8	2.6	2.9
Camera	3.6	3.0	3.4
Cane	4.1	4.9	4.0
Cannon	3.3	2.2	3.1
Car	2.9	2.4	2.8
Carrot	3.6	4.9	3.8
Cat	2.8	3.3	2.9
Chair	4.1	4.6	4.1
Cheese	3.9	2.7	3.5
Clock	2.8	3.1	2.7
Comb	3.7	2.5	3.3
Corn	3.6	3.5	3.3
Couch	3.2	2.4	2.8
Dice	3.3	3.5	3.4
Doll	3.5	4.8	3.6
Earrings	2.7	3.7	3.0
Eggs	2.8	2.7	3.0
Elephant	3.9	4.7	3.8
Fan	3.9	4.8	4.1
Feather	3.9	4.7	4.0

Fireplace	3.6	4.5	3.6
Fish	3.0	4.8	3.1
Flashlight	3.0	2.4	3.1
Flower	2.8	4.4	2.9
Football	3.4	4.7	3.6
Fork	3.8	4.8	3.8
Frog	3.7	4.8	3.8
Glasses	4.0	4.7	4.0
Globe	3.7	4.4	3.9
Glove	3.6	4.7	3.7
Guitar	3.1	3.7	3.1
Hammer	3.1	2.3	2.9
Helicopter	3.4	2.1	2.9
Horse	3.4	3.2	3.2
Horse shoe	2.6	4.3	2.9
Iron	3.3	2.2	3.0
Key	2.8	4.9	3.0
Ladder	3.5	2.6	3.6
Lamp	2.7	4.6	3.1
Leaf	2.4	4.3	2.8
Lighter	2.7	4.8	3.1
Lion	4.3	2.3	3.7
Lipstick	3.2	4.5	3.3
Lock	2.6	3.7	2.7
Mailbox	2.8	4.1	3.0
Medal	4.1	4.9	4.0
Motorcycle	3.5	2.4	3.2
Mouse	3.0	2.5	3.0
Muffin	3.5	4.6	3.4
Mushroom	3.1	4.7	3.3
Notebook	3.6	3.1	3.4
Owl	2.6	4.0	2.8
Pacifier	3.6	2.3	3.3
Pear	3.8	4.7	3.9
Pen	3.2	4.8	3.4
Penguin	3.5	2.9	3.4
Piano	3.6	4.7	3.7
Pie	3.0	3.6	3.3
Pill	2.8	3.6	2.7
Pineapple	3.7	4.7	3.6
Pipe	3.7	4.9	3.7
Potato	3.7	4.7	3.6
Pretzels	4.5	3.1	4.1
Purse	2.9	2.7	2.9
Roller blade	3.3	4.7	3.6
Ruler	3.7	4.7	3.7

Saxophone	3.3	2.7	3.2
Scissors	3.1	3.8	3.2
Shark	3.6	2.3	3.3
Shovel	3.2	2.9	3.3
Skateboard	2.9	4.3	3.2
Sponge	2.5	3.9	2.8
Spoon	3.3	4.9	3.5
Stapler	2.5	2.4	2.9
Swing set	3.4	4.2	3.4
Tent	3.1	4.3	3.3
Tie	2.9	4.0	3.0
Tire	3.1	4.8	3.2
Toaster	2.9	3.1	2.9
Toilet	3.3	2.3	3.1
Tomato	4.0	4.5	4.0
Tractor	2.6	2.4	2.7
Trophy	2.7	4.6	2.8
Turtle	3.1	4.5	3.4
Tweezers	2.8	2.4	2.6
Typewriter	2.8	2.9	2.8
Umbrella	3.2	3.6	3.1
Van	3.1	4.3	3.5
Violin	3.5	2.4	3.2
Watch	3.4	2.3	3.1
Watermelon	3.3	2.9	2.8
Wheelchair	3.4	2.4	3.3
Whistle	3.0	4.3	3.3
Wreath	3.5	4.8	3.5
Yarn	4.0	4.7	3.9
Yoyo	3.7	2.5	3.4
Zebra	4.3	2.2	3.7

APPENDIX F

Method for Silhouette Study

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

Stimuli. The stimuli were the target objects used in Experiment 1, silhouettes of the targets created by reducing the contrast in the original objects until the interior was blackened, and the non-object control.

Apparatus and Procedure. The apparatus and procedure were identical to Experiment 2, except that there were only three preview conditions (same, silhouette, and the control) and only 12 practice trials.

Results

As in Experiments 1 and 2, mean naming latencies excluded trials in which the target object was named incorrectly, an anticipatory eye movement occurred (saccade latencies of less than 100 ms), and trials on which the naming latency was less than 200 ms or more than 3 standard deviations greater than the mean naming latency for that subject. Eliminated trials accounted for 8% of the data.

Naming latencies were subjected to a within-participants ANOVA, which revealed reliable differences across the preview conditions, $F(2,40) = 14.306$, $MSE = 3867$, $p < .001$. Naming latencies were 102 ms faster in the identical preview condition (mean = 796 ms) than in the control condition (mean = 898 ms), $F(1,20) = 28.574$, $MSE = 3844$, $p < .001$, and the 43 ms faster in the silhouette condition (mean = 855 ms) than in the control condition, $F(1,20) = 4.382$, $MSE = 4559$, $p < .05$. The 59 ms advantage for the

identical preview over the silhouette was also reliable, $F(1,20) = 11.302$, $MSE = 3197$, $p < .01$.

References

- Balota, D. A., & Rayner, K. (1983). Parafoveal visual information and semantic contextual constraints. *Journal of Experimental Psychology: Human Perception and Performance*, 9, 726-738.
- Biederman, I. (1987). Recognition-by-components: A theory of human image understanding. *Psychological Review*, 94, 115-147.
- Breitmeyer, B. G., Kropfl, W., & Julesz, B. (1982). The existence and role of retinotopic and spatiotopic forms of visual persistence. *Acta Psychologica*, 52, 175-196.
- Cutzu, F., & Tarr, M. J. (1997). The representation of three-dimensional object similarity in human vision. In *SPIE proceedings from electronic imaging: Human vision and electronic imaging ii* (Vol. 3016, p. 460-471). San Jose, CA: SPIE.
- Davidson, M. L., Fox, M. J., & Dick, A. O. (1973) Effect of eye movements on backward masking and perceived location. *Perception & Psychophysics*, 14, 110-116.
- Eimas, P. D., & Quinn, P. C. (1994). Studies on the formation of perceptually based basic-level categories in young infants. *Child Development*, 65, 903-917.
- Feldman, J. A. (1985). Four frames suffice: A provisional model of vision and space. *Behavioral and Brain Sciences*, 8, 265-289.
- Gajewski, D. A., & Henderson, J. M. (2005). The role of saccade targeting in the transsaccadic integration of types and tokens. *Journal of Experimental Psychology: Human Perception and Performance*, 31, 820-830.
- Hayward, W. G. (1998). Effects of outline shape in object recognition. *Journal of Experimental Psychology: Human Perception and Performance*, 24, 427-440.
- Helmholz (1867, 1925). Treatise on physiological optics. Ed and trans. J. P. C. Southall. New York: Optical Society of America.
- Henderson, J. M. (1992a). Identifying objects across eye fixations: Effects of extrafoveal preview and flanker object context. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 18, 521-530.
- Henderson, J. M. (1992b). Visual attention and eye movement control during reading and picture viewing. In K. Rayner (Ed.), *Eye movements and visual cognition: Scene perception and reading* (pp. 260-283). New York: Springer-Verlag.
- Henderson, J. M. (1994). Two representational systems in dynamic visual identification. *Journal of Experimental Psychology: General*, 123, 410-426.

- Henderson, J. M. (1997). Transsaccadic memory and integration during real-world object perception. *Psychological Science*, 8, 51-55.
- Henderson, J. M., & Anes, M. D. (1994). Roles of object-file review and type priming in visual identification within and across eye fixation. *Journal of Experimental Psychology: Human Perception and Performance*, 20, 826-839.
- Henderson, J. M., Pollatsek, A., & Rayner, K. (1987). The effects of foveal priming and extrafoveal preview on object identification. *Journal of Experimental Psychology: Human Perception and Performance*, 13, 449-463.
- Henderson, J. M., Pollatsek, A., & Rayner, K. (1989). Covert visual attention and extrafoveal information use during object identification. *Perception & Psychophysics*, 45, 196-208.
- Henderson, J. M., & Siefert, A. B. C. (2001). Types and tokens in transsaccadic object identification: Effects of spatial position and left-right orientation. *Psychonomic Bulletin & Review*, 8, 753-760.
- Hoffman, J. E., & Subramaniam, B. (1995). The role of visual attention in saccadic eye movements. *Perception & Psychophysics*, 57, 787-795.
- Irwin, D. E. (1993). Perceiving an integrated visual world. In D. E. Meyer and S. Kornblum (Eds.), *Attention and performance: Synergies in experimental psychology, artificial intelligence, and cognitive neuroscience- A silver jubilee* (pp. 121-142). Cambridge, MA: MIT Press.
- Irwin, D. E., Brown, J. S., & Sun, J. S. (1988). Visual masking and visual integration across saccadic eye movements. *Journal of Experimental Psychology: General*, 117, 276-287.
- Irwin, D. E., Yantis, S., & Jonides, J. (1983). Evidence against visual integration across Saccadic eye movements. *Perception & Psychophysics*, 34, 49-57.
- Irwin, D. E., Zacks, J. L., & Brown, J. S. (1990). Visual memory and the perception of a stable visual environment. *Perception & Psychophysics*, 47, 35-46.
- Johnson, C. J., Paivio, A., & Clark, J. M. (1996). Cognitive components of picture naming. *Psychological Bulletin*, 120, 113-139.
- Jonides, J., Irwin, D. E., and Yantis, S. (1982). Integrating visual information from successive fixations. *Science*, 215, 192-194.
- Kahneman, D., & Treisman, A. (1984). Changing views of attention and automaticity. In R. Parasuraman & R. Davies (Eds.) *Varieties of attention* (pp. 29-61). Cambridge, MA: MIT Press.

- Kahneman, D., Treisman, A., & Gibbs, B. J. (1992). The reviewing of object files: Object-specific integration of information. *Cognitive Psychology*, 24, 175-219.
- Kanwisher, N., & Driver, J. (1992). Objects, attributes, and visual attention: Which, what, and where. *Current Directions in Psychological Science*, 1, 26-31.
- Kowler, E., Anderson, E., Doshier, B., & Blaser, E. (1995). The role of attention in the programming of saccades. *Vision Research*, 35, 1897-1916.
- Matin, E. (1974) Saccadic suppression: A review and an analysis. *Psychological Bulletin*, 81, 899-917.
- McClelland, J. L., & O'Regan, J. K. (1981). Expectations increase the benefit derived from parafoveal visual information in reading words aloud. *Journal of Experimental Psychology: Human Perception and Performance*, 7, 634-644.
- McConkie, G. W., & Rayner, K. (1976). Identifying the span of the effective stimulus in reading: Literature review and theories of reading, In H. Singer & R. B. Ruddell (Eds.), *Theoretical models and processes of reading* (pp. 137-162). Newark, NJ: International Reading Association.
- McConkie, G. W., & Zola, D. (1979). Is visual information integrated across successive fixations in reading? *Perception & Psychophysics*, 25, 221-224.
- Neisser, U. (1967). *Cognitive psychology*. Englewood Cliffs, NJ: Prentice-Hall.
- O'Regan, J. K., & Lévy-Schoen, A. (1983). Integrating visual information from fixations: Does trans-saccadic fusion exist? *Vision Research*, 23, 765-768.
- Paap, K. R., & Newsome, S. L. (1981). Parafoveal information is not sufficient to produce semantic or visual priming. *Perception & Psychophysics*, 29, 457-466.
- Palmer, S. E., Rosch, E., & Chase, P. (1981). Canonical perspective and the perception of objects. In J. Long & A. Baddeley (Eds.), *Attention and Performance* (Vol. 9) (pp. 135-151). Hillsdale, NJ: Erlbaum.
- Poggio, T., & Edelman, S. (1990). A network that learns to recognize three-dimensional objects. *Nature*, 343, 263-266.
- Pollatsek, A., Lesch, M., Morris, R. K., & Rayner, K. (1992). Phonological codes are used in integrating information across saccades in word identification and reading. *Journal of Experimental Psychology: Human Perception and Performance*, 18, 148-162.
- Pollatsek, A., Rayner, K., & Collins, W. E. (1984). Integrating pictorial information

- across eye movements. *Journal of Experimental Psychology: General*, 113, 426-442.
- Pollatsek, A., Rayner, K., & Henderson, J. M. (1990). Role of spatial location in integration of pictorial information across saccades. *Journal of Experimental Psychology: Human Perception and Performance*, 16, 199-210.
- Quinn, P. C., Eimas, P. D., & Tarr, M. J. (2001). Perceptual categorization of cat and dog silhouettes by 3- to 4-month-old infants. *Journal of Experimental Child Psychology*, 79, 78-94.
- Rayner, K. (1975). The perceptual span and peripheral cues in reading. *Cognitive Psychology*, 7, 65-81.
- Rayner, K. (1978). Foveal and parafoveal cues in reading. In J. Requin (Ed.), *Attention and performance* (Vol. 7, pp. 149-162). Hillsdale, NJ: Erlbaum.
- Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, 85, 618-660.
- Rayner, K., McConkie, G. W., & Ehrlich, S. (1978). Eye movements and integrating information across fixations. *Journal of Experimental Psychology: Human Perception and Performance*, 4, 529-544.
- Rayner, K., McConkie, G. W., & Zola, D. (1980). Integrating information across eye movements. *Cognitive Psychology*, 12, 202-226.
- Rayner, K., & Pollatsek, A. (1983). Is visual information integrated across saccades? *Perception & Psychophysics*, 34, 39-48.
- Riesenhuber, M., & Poggio, T. (2000). Models of object recognition. *Nature Neuroscience*, 3, 1199-1204.
- Ritter, M. (1976). Evidence for visual persistence during saccadic eye movements. *Psychological Research*, 39, 67-85.
- Shepherd, M., Findlay, J. M., & Hockey, R. J. (1986). The relationship between eye movements and spatial attention. *The Quarterly Journal of Experimental Psychology*, 38A, 475-491.
- Tarr, M. J., & Bülthoff, H. H. (1998). Image-based object recognition in man, monkey, and machine. *Cognition*, 67, 1-20.
- Tarr, M. J. (2003). Visual object recognition: Can a single mechanism suffice? In M. A. Peterson & G. Rhodes (Eds.), *Perception of faces, objects, and scenes: Analytic and holistic processes* (pp. 177-211). Oxford, UK: Oxford University Press.

- Tarr, M. J., Williams, P., Hayward, W. G., & Gauthier, I. (1998). Three-dimensional object recognition is viewpoint-dependent. *Nature Neuroscience*, 1, 275-277.
- Treuhub, A. (1977). Neuronal models for cognitive processes: Networks for learning, perception, and imagination. *Journal of Theoretical Biology*, 65, 141-169.
- Treisman, A. (1993). Representing visual objects. In D. E. Meyer & S. Kornblum (Eds.), *Attention and Performance XIV: Synergies in experimental psychology, artificial intelligence, and cognitive neuroscience* (pp. 163-175). Cambridge, MA: MIT Press.
- Ullman, S. (1998). Three-dimensional object recognition based on the combination of views. *Cognition*, 67, 21-44.
- 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* (pp. 549-586). Cambridge, MA: MIT Press.
- Wolf, W., Hauske, G., & Lupp, U. (1978) How pre-saccadic gratings modify post-saccadic modulation transfer functions. *Vision Research*, 18, 1173-1179.
- Wolf, W., Hauske, G., & Lupp, U. (1980). Interactions of pre- and post-saccadic patterns having the same coordinates in space. *Vision Research*, 20, 117-125.