THE PERCEPTUAL SPECIFICITY OF REPRESENTATIONS OF COMPLEX SCENES FOR ROTATION, TRANSLATION, AND REFLECTION

BY

MARK WILLIAM OLSON

B.A., Taylor University, 1990
A.M., University of Illinois, 1992

THESIS

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Neuroscience
in the Graduate College of the
University of Illinois at Urbana-Champaign, 1998

Urbana, Illinois

Abstract

The issue of how complex visual scenes are represented has received relatively little attention. Perceptual specificity refers to the degree to which a physical change to a stimulus, upon re-presentation, alters performance in any type of memory test. I investigated the perceptual specificity of scene representations for various kinds of changes to naturalistic scenes using eye movements and a direct measure of memory (familiarity or recognition). In Experiment 1, both measures indicated that scene representations, like object representations, are perceptually specific for rotations in depth. In Experiment 2, I replicated this result using backgroundless outdoor objects and concluded that this similar result emerged because the representations of both objects and scenes have view dependent components, and because the primary objects in a scene are allotted more weight in the scene's representation. In Experiment 3, I horizontally translated scenes in order to determine if scene representations are frame-specific and to test the hypothesis that primary objects are allotted more representational weight. I found that scene representations are senstive to frame translation in a manner that suggests that primary objects are weighted more heavily. In Experiment 4, I altered the left-right orientation of the picture to determine whether sensitivity for translation in Experiment 3 was due to deletion of scene parts or changes in the spatial relationships between scene features and the viewer. Mild specificity for orientation suggests that both play a role. I suggest that the perceptual specificity of these results in all 4 experiments supports the view that the representations of complex scenes have components which are view-, frame-, and orientation-dependent. Future research is suggested for examining further the nature of complex scene representations.

Dedication

To Stacy and Cindy, two very beautiful people.

Acknowledgments

I would like to thank the numerous people who played a significant and necessary part of this thesis:

Art Kramer, for being extremely available and for generously providing significant amounts of support and tons of thorough advice during the latter half of this project.

Neal Cohen, for providing me with the awesome invaluable opportunity to address these issues with these means, and for clarifying the graduate school path and supporting me during the first half of this project.

All my committee members (Art Kramer, Neal Cohen, Marie Banich, George McConkie, David Irwin, Tom Anastatio, and Matilde Holzwarth) for offering their time and insights during the course of this project.

Rob Althoff, for teaching me how to use Adobe Photoshop, the eye tracker, and numerous other computer skills when I desperately needed it.

Mike Maciukenas, Chris Lopez, Rob Althoff, Tim Weber, Roger Marsh, the Viz Lab staff, and the BISS staff, for tons of much-needed technical assistance and for putting up with my lack of computer skills.

Razia Azen, for help with statistics and statistical programs.

Don, Peter, Brian, Tamara, Liz, Michele, Amanda, and Jen, for assistance preparing stimuli, writing programs, running subjects, and organizing files.

Mom and Dad, for an effective combination of genetics and environment to get me here and

the support to get me out.

Ralph Page, for unknowingly serving as my role-model academian.

Stacy, Cindy, and Jane, for the tremendous social and emotional support that I could not have done without and will never forget.

Many thanks to all of you—this thesis exists because you do!

Table of Contents

List of Tables

List of Figures

Introduction

While viewing a photo album with someone, the person unfamiliar with the places depicted in the photos is often unable to notice when a place is depicted again. After continued exposure, however, the person may be able to recognize “your house” even when viewing a new picture of it. How much exposure and what kind of exposure must one have until one is capable of recognizing any new instance of a place? What kinds of changes between pictures can occur without recognition being impaired? Would your house be recognized by someone else if the view changed to a close-up or to a different side of the house, or if the shrubs in front of the house or the trees in back of it were removed, or if the second picture is “shifted” horizontally from the first, or the picture is flipped horizontally like a slide? Might that individual, while not expressing any conscious recognition for new pictures of your house, nevertheless report a sense of familiarity while viewing these pictures or view them differently than he/she otherwise would have had he/she not seen any pictures in the first place?

These are questions that address the issue of how scenes are represented in our memory. Specifically, these questions address the perceptual specificity of scene representations, or the degree to which a stimulus, upon re-presentation, must resemble its form at initial presentation for performance on any one of a number of various memory tasks not to be affected. If a particular measure shows high perceptual specificity for a particular kind of stimulus attribute, such as viewpoint, this means that the more the view of the object at test departs from the view of the object at study, the more performance in this measure will be impaired. While determining the perceptual specificity of a measure is interesting in and of itself, ultimately we are interested in making inferences about how the brain represents the external world. In other words, the degree to which performance on these tasks is perceptually specific or nonspecific informs us about the nature of the representation of the stimulus being viewed, so by determining the perceptual specificity of a particular measure, we can make inferences about the representation that underlies performance in this measure. For example, if perceptual specificity is found in relationship to a particular dimension, such as viewing angle, we might infer that the representation for the object is view dependent (with respect to rotation), or at least possesses view dependent components. If changing the view of an object between presentations has no effect on performance, however, we might infer that the representation of the object is object-centered, although this need not be the case as I will explain later.

In this thesis, I investigate the nature of the representations for pictures of outdoor scenes. Ultimately, we endeavor to determine what types of representations exist for the real, three-dimensional scenes that we experience, so I use two-dimensional color photographs of everyday scenes as stimuli because it is a step closer to examining three-dimensional scenes than most research that uses pictorial stimuli consisting mainly of line-drawings or photographs of single backgroundless objects.

In order to better understand why complex scene representation is an interesting issue, it is helpful to first consider how objects are represented. Both phenomenal experience and research utilizing electrophysiological recordings and brain imaging suggest that objects are represented “as objects.” By this I mean that visual processing does not end with feature detection, but rather continues forward to construct, from subordinate features, superordinate entities known as objects (van Lier, Leeuwenberg, van der Helm, 1997). Phenomenally, we are aware of objects around us and often must focus our attention to even become aware of an object’s constituent features. Electrophysiological research has revealed that cells in the fusiform gyrus and inferior temporal lobe (IT) fire preferentially to particular objects, such as faces, but not preferentially for a face’s features in a scrambled arrangement (Rolls, Tovee, Purcell, Stewart, & Azzopardi, 1994). Brain imaging studies (Puce, Allison, Gore, & McCarthy, 1995) also find that specific areas of the temporal lobe are more active when normal faces are displayed rather than the rearranged parts of faces. In contrast, research by Wachsmuth, Oram, and Perrett (1997) has found neurons which fire preferentially for specific body parts, but not the body as a whole. So it seems that our brains often, though not always, represent particular spatial arrangements of features as (superordinate) objects, not just as (subordinate) features.

This leads us to the issue of scenes, which consist of spatial relationships between objects and hence in some ways seem to resemble a superordinate “object” constructed from subordinate objects. Do scenes also enjoy a special status as higher level “objects,” with objects as “features,” or do we reach maximum complexity at the level of objects? Clearly there are differences between scenes and objects. One cannot readily point to a scene or its boundaries as one can an object, and while objects are generally opaque, allowing us only a view of their external surfaces, scenes are transparent in object terms. Nevertheless, we do refer to scenes verbally and the encoding of spatial relationships between objects seems analogous to the encoding of spatial relationships between features. If we decide that scenes are represented as unique "objects," then we must also determine whether this representation is best characterized as being view-dependent or something more abstract and thereby view invariant (or both). These are some of the higher level issues that this thesis begins to address but by no means resolves.

On a simpler level, this thesis investigates the perceptual specificity of scene representations for rotation, translation, and reflection. Each of these types of changes manipulates a particular aspect of the picture in relation to the viewer that allows us to partially ascertain the attributes of the scene representation. In Experiment 1, where the scene rotates in depth, the picture changes because every object rotates with respect to the viewer while the spatial relationships between the objects remain the same. In Experiment 2, the same rotations are applied to outdoor objects with no backgrounds, thus removing the component of the spatial relationships between objects and leaving only the rotation of the center object as the source of effects. In Experiment 3, where we horizontally shift the picture relative to its frame, the picture as a whole changes, like the other experiments, but no objects rotate with respect to the viewer; thus, the scene as a whole changes even though the primary object does not change. In Experiment 4, the picture is reflected, thus producing a change in the two-dimensional picture but not a change in the three-dimensional scene, the objects within the scene, or the spatial relationships between them. To summarize, in Experiment 1, the picture, the scene, and the objects change. In Experiment 2, the picture and object changes, but the scene does not. In Experiment 3, the picture and the scene changes but the primary object does not. In Experiment 4, the picture changes, but the scene and the objects do not. In none of these experiments do spatial relationships of any kind change between objects. These experiments serve as approximations for the combinations of changes and nonchanges for scene-object-viewer relations that we are trying to construct. More controlled versions of these combinations, particularly for Experiments 2 and 3, are discussed in the Future Directions section.

Object and Scene Representation

In this section, I review the previous research that directly or indirectly addresses the issue of representation of objects and scenes, particularly the research that addresses perceptual specificity of representations for changes in rotation, translation, and reflection of a stimulus. Most of the research investigating the perceptual specificity of representations for rotation in depth have used single objects (line drawings and photgraphs) as stimuli. The extent to which such research can be applied to understanding scene representation is not clear since, as noted before, scenes are composed of multiple objects with spatial relationships between them and it is also unclear whether these spatial relationships are coded in a qualitatively similar way to the kinds of spatial relationships that exist between parts of an object. Given these potential differences, I discuss the object literature separately from the scene literature, which includes research conducted on spatial navigation and wayfinding.

Rotation of Objects. A significant amount of research has been conducted on the topic of visual object representation and much of it addresses the issue of perceptual specificity for object rotation (Logothetis & Sheinberg, 1996). The general finding that will be discussed in greater detail below is that rotation of objects in the depth plane (around the y axis) results in a decrement of recognition or identification performance relative to the performance when the object is not rotated. Most theorists suggest that such results lend support for the existence of viewer-centered (or view-dependent) representations (Logothetis, Pauls, Bulthoff, & Poggio 1994; Vetter, Hurlbert, & Poggio 1995; Bulthoff & Edelman, 1992), but some argue that these results can be accounted for by object-centered representations alone (Biederman & Gerhardstein, 1993), though their claims are rebutted by Tarr, Bultoff, Zabinski, and Blanz (1997).

A significant debate exists between those who believe that objects are represented in a viewer-centered manner and those who believe that objects are represented in an object-centered manner (Ullman, 1989). Viewer-centered theories suggest that objects are represented from more than one viewpoint; recognition of an object from a novel view requires the matching of the image with the representation by means of a rotation process. The images that are represented are presumably those for which the individual has had the most exposure, and hence it is thought that these representations are canonical, by definition. Viewer-centered representations are normally equated with perceptually specific performance on a task, but if enough viewer-centered representations exist, or if these representations are very broadly tuned, perceptually nonspecific performance on a task may emerge. Furthermore, while perceptually nonspecific performance is normally equated with object-centered representations, perceptual changes that result in the occlusion of object parts can lead to perceptually specific performance on a task. Thus, if amount of exposure and part occlusion is not controlled for properly, there need not be a direct relationship between perceptual specificity of a measure and the type of representation underlying performance on that measure, since both perceptually specific and nonspecific effects can emerge from either object-centered or viewer-centered representations. Furthermore, the type of task that the subject engages in may rely on different types of representations. For example, Srinivas (1995) suggests that what underlies the lack of perceptually specific effects in her implicit measures is a view-invariant (or object-centered) representation of the object, and that the perceptually specific effects that emerged in her explicit measures arise from a view-dependent (or viewer-centered) representation of the object. In other words, different types of representations were accessed for different types of tasks.

Support for the existence of view dependent representations comes from both behavioral and electrophysiological data (Tanaka, 1993). Rotation of objects in the picture plane or in depth result in both reduced priming and recognition for objects (Bartram, 1974; Srinivas, 1993, 1995; Tarr, 1989; Tarr, Bulthoff, Zabinski, & Blanz, 1997). Picture matching tasks in which rotated line drawings of objects are not displayed simultaneously have shown a direct relationship between view rotation and reaction time (Lawson & Humphreys, 1996).

In addition, single cell recordings in primate extrastriate regions V4 and IT reveal cells that are view-dependent in their firing characteristics for objects (Logothetis & Pauls, 1995; Logothetis, et al., 1994, 1995; Schiller, 1995) and faces (Perrett, Smith, Potter, Mistlin, Head, Milner, & Jeeves, 1985). Kendrick, Atkins, Hinton, Broad, Fabre-Nys, and Keverne (1995) have found similar results in sheep. Computational models of recognition have been developed that mimic these behavioral and electrophysiological findings and lend support to the viability of the viewer-centered theories (Bulthoff & Edelman, 1992; Edelman & Weinshall, 1991; Poggio & Edelman, 1990; Vetter, et al., 1995).

Object-centered theories claim that objects are represented in a manner that is not view-dependent, but rather view-invariant. This representation is thought of by some as a structural description (Hummel & Biederman, 1992) that depends on the ability for one to discern the parts of an object (Biederman & Gerhardstein, 1993, 1995) while others have suggested that these representations are concerned with an object's axes of elongation (Marr, 1982). In either case, recognition does not require rotation, but rather matching the on-line structure with the structure represented in memory (Quinlan, 1991). To explain the view-dependent effects that have been found in the priming studies mentioned above, Biederman & Gerhardstein (1993, 1995) suggest that the designs of these experiments confound rotation with occlusion of critical object parts. They suggest that priming is based on object-part descriptions; thus, as long as the same parts of the object are visible and distinct from the parts of other objects, priming should be view-invariant across an experiment (but see Srinivas, 1995). Object-centered representations are perceptually nonspecific as long as the critical attributes of the stimulus, whatever they may be, are discernable. In support of object-centered theories, electrophysiological studies (Logothetis, Pauls, & Poggio, 1995) have revealed cells that fire in a manner that is view-invariant, or object-centered, though it is unknown whether these cells fire as a result of convergent inputs from view-dependent cells or whether they fire without such input, strictly on the basis of the perceived structure or axes of the stimulus.

Rotation of Scenes. The previous section described the research performed for the purpose of determining how objects are represented. The focus of this thesis, however, is how complex scenes are represented. Very little research has been conducted in this area and it is not entirely clear the extent to which the object representation literature applies to scene representation. Are scenes represented as higher-level objects (superordinate "objects" composed of subordinate objects), or is it the case that there exists no scene representations, only object representations? Are scenes represented in a view dependent manner or in a more abstract manner analogous to object-centered representations?

Research on spatial navigation and wayfinding has contributed to our understanding of how scenes are described and represented in memory. Many researchers have found that scenes are described (and presumably represented) in a variety of ways, and Taylor and Tversky (1996) find that the method used to describe and represent a scene depends in part on the type of scene in question. For instance, for small environments that can be described from a single viewpoint (such as the one used in the Diwadkar and McNamara (1997) study described below), people describe the scene by describing the spatial relationships between objects from an outside viewer-centered perspective. This type of description is called a "gaze tour." When the scene is too large to be described from a single viewpoint, such as the scenes used in this thesis, one of two additional methods are used: a route perspective or a survey perspective. A route perspective, like a gaze tour, is viewer-centered in its reference frame, but it is different from a gaze tour in that the viewer's perspective keeps changing. In a survey perspective, an external reference frame is used that is based on the canonical direction of scene landmarks, and the scene is described as if from above or like a map. Thoryndyke and Hayes-Roth (1982) claim that at least two types of survey knowledge exist. Both are highly abstract from the viewer’s perspective, but one is best described as having a map-like perspective while the other allows the viewer to “see through” obstacles to other objects in the environment.

Other researchers have made similar trichotomies to the one proposed above by Taylor and Tversky (1996). For instance, the distinction between a deictic or viewer-centered perspective, an intrinsic or object-centered perspective, and an extrinsic or environment-centered perspective has been suggested by a number of researchers (Carlson-Radvansky & Irwin, 1994; Levelt, 1989; Shepard & Hurwitz, 1984).

Relating these conceptions to the object recognition literature, the gaze tour or deictic perspective is most analogous to a viewer-centered representation since both are intimately tied to the viewer’s perspective. The survey or environment-centered perspective is most analogous to the object-centered perspective since they are both representations that are abstracted from the viewer’s experience. One critical difference between a survey perspective (for scenes) and an object-centered representation (for objects) is that the latter can be formed from a single exposure while the former requires having been made familiar with the scene via real or imagined “tours” of the scene or by studying a map of the scene. It should also be noted that an object-centered perspective in the scene literature is not analogous to an object-centered representation in the object recognition literature; with scenes, an object-centered perspective refers to a perspective in which the environment is described in terms of the canonical reference frame of an object within the scene and in terms of the spatial relations of other objects with that object. For the purpose of this thesis, we are concerned about whether scenes are represented in a more experience-based, view-dependent manner or in a more abstract manner, similar to a survey, or environment-based, perspective.

Diwadkar and McNamara (1997) conducted a study in which subjects studied scenes of small indoor objects and were asked to recognize each scene from different viewpoints. Each scene was composed of the same 6 objects, and only the spatial relationships between the objects changed between scenes. They found that recognition latency was a linear function of the angular distance between the test view and the nearest study view and suggest that between-object spatial relationships are represented in a view-dependent manner. In addition they stress that since all of their scenes were composed of the same 6 objects in different spatial arrangements, that scene representations are composed of more than just the objects in the scene; rather, scene representations include the between-object spatial relationships as well. Similarly, Shelton and McNamara (1997) conducted a study in which subjects were placed in a large room with numerous objects and were later asked to point to a particular object while they imagined that they were standing in a specified place in the room. They found that subjects were better able to perform this task when the imagined position was aligned with the original learned view rather than from a new view that was not aligned. Such results, again, support the idea that scenes are represented in a view-dependent manner.

Translation of Objects. Horizontal translation of a scene is another type of manipulation that can be used to explore the nature of scene representations. The type of scene translation that is performed in the present research has not been performed by any other researchers. The most similar research is that by Biederman and Cooper (1991, 1992) in which they translate objects horizontally relative to retinal position between study and test and ask subjects to identify the object. This experiment was performed once using line drawings and then again using photographs and in both cases they found that object identification was invariant to horizontal translation. In these same experiments, they also found that object identification was invariant to changes in size. They concluded that the representations mediating performance on their task were invariant to changes in horizontal retinal position and size. The relevance of this research to the present study, however, is unknown since the Biederman and Cooper studies were priming studies which shifted position relative to the subject's retina while the present study is a recognition study that shifts position relative to the frame of the picture.

Translation of Scenes. The only research performed on the translation of scenes is the work of Intraub and colleagues (Intraub & Richardson, 1989; Intraub & Berkowits, 1996; Intraub, Gottesman, Willey, & Zuk, 1996). In this research, subjects are shown a photograph of an outdoor scene and are later tested for their memory of that scene. Intraub consistently finds that subjects recall scenes as having wider boundaries than they actually had. This effect emerged whether subjects were asked to draw what they recall or choose from a set of options the picture they saw before. Intraub argues that this effect is due to a schematization of the scene where objects that would be expected to be in the scene are recalled as having been there, just as similar work by Brewer and Treyens (1981) found that subjects would recall having seen books on an office shelf even though no books were present in the studied photograph. In terms of this study, Intraub found that scene representations are not specific with respect to picture frame.

In the present research, we investigate this issue further by translating the frame horizontally, thereby widening the frame on one side by narrowing it on the other. By doing this, we can determine whether the representations are simply not frame-specific or whether they are truly wider than the study image. If subjects cannot detect a frame shift, then we would conclude that our effects and Intraub's effects may both be due to scene representations simply not being frame-specific. If, however, subjects do detect the shift, then we can conclude that scene representations may encompass more than the picture depicts.

Reflection of Objects. Research on object reflection has often found a dissociation between perceptually specific effects using recognition memory measures and nonspecific effects in priming paradigms. For instance, Seamon, Ganor-Stern, Crowley, Wilson, Weber, O-Rourke, and Mahoney (1997) found that recognition memory for line drawings of objects was not invariant to reflection while ratings of affective preference (the mere exposure effect) were invariant to reflection. Cooper, Schacter, Ballesteros, and Moore (1992) and Zimmer (1995) have also found similar results in which recognition memory (but not priming) is sensitive to changes in left-right orientation. However, Martin and Jones (1997) found that recognition memory for the moon's orientation was very poor, suggesting that object representations are not orientation-dependent. Similar effects are found when subjects are asked to recall the direction in which the head of a coin is facing; performance on this task is very poor (Jones, 1990; Martin & Jones, 1995).

Since we are interested in the representations that mediate recognition memory, studies involving priming or the mere exposure effect are not relevant here; the moon orientation study may be a special case where recognition is not sensitive to orientation since we do not interact with celestial bodies motorically. For similar reasons, coin studies may be irrelevant as well since the orientation of the head on a coin has no motoric value for us either. For these reasons, it seems best to focus on the Seamon, et al., (1997) study and tentatively conclude that object representations are orientation-dependent.

Reflection of Scenes. The research on scene reflection shows a similar pattern of dissociative results in which priming studies show orientation-invariant results while recognition memory studies show orientation-dependent results. Bartlett, Gernsbacher, and Till (1987) asked subjects to make “same/different” decisions concerning the orientation of previously studied photographs of landscapes and cityscapes. They found that accurate classification was higher for identical pictures than reflected ones and offered a dual process model to explain these results. These results are consistent with previous research using similar paradigms with faces as stimuli (Gernsbacher, 1985; Klatsky & Forrest, 1984; McKelvie, 1983). In terms of the present study, these findings suggest that both orientation-dependent as well as orientation-invariant scene representations exist. Had they found that both “same” and “different” decisions were equally accurate, this would have suggested that representations are only orientation-dependent, but the lower accuracy for the reflected images suggests that orientation-invariant representations were also mediating responses. In other words, since accuracy was lower for reflected images, this indicates that (orientation-invariant) representations that were not sensitive to the scene being in a different orientation were mediating the subjects’ responses. But because performance was not at chance for reflected images, orientation-dependent representations had to also be mediating performance.

Other research on scene orientation has required subjects to classify both nonreflected and reflected scenes as “old,” rather than having subjects distinguish between the two. This is the paradigm that was used in Experiment 4 of the present thesis. This previous research has often found a slightly more accurate recognition performance for the identical scenes over the reflected scenes (Bartlett, Till, Gernsbacher, & Norman, 1983; Bartlett, Till, & Levy, 1980), though this effect is not always present (Standing, Conezio, & Haber, 1970). Those studies that have found this difference support the argument that orientation-dependent scene representations exist. If no difference was found with this paradigm, one would argue for orientation-invariant representations, but the difference requires the presence of orientation-dependent representations.

In another experiment by Frederickson and Bartlett (1987), subjects learned the layout of a scene that included in its center an image projected onto a screen with a slide projector. Half of the subjects viewed the whole scene directly while the other half viewed the entire scene through a mirror. At test, the image projected on the screen either remained the same or was reflected and subjects either viewed the entire scene directly or through a mirror such that all possible orientation combinations between study and test were counterbalanced. Subjects were then asked either to state whether the projected scene was the same relative to them (viewer-centered) or were asked to state whether the projected scene was the same relative to the rest of the room (scene-centered). Subjects' performances were better when asked the viewer-centered question than the environment-centered question, suggesting that spatial relationships between objects are represented in a viewer-centered manner, rather than a scene-centered manner.

In summary, the "same/different" and "old/old" recognition studies on scene orientation both show support for the argument that both orientation-invariant and orientation-dependent representations exist. The research by Frederickson and Bartlett supports the existence of orientation-dependent (viewer-centered) representations.

Eye Movements

Eye movements can provide a rich set of measures on which to infer the kinds of cognitive activity in which an individual is engaged. To understand the kind of information eye movements can provide and how eye movements can be used to study prior exposure with a stimulus, a brief overview of eye movement research follows.

While viewing a stationary picture, a person's eyes make an alternating pattern of fixations, during which the eyes are stationary and information is extracted, and saccades, during which the eyes move between fixation locations. Previous eye movement research has attempted to understand what determines the spatio-temporal pattern of eye movements and has found that many factors contribute. Buswell (1935) found that subjects fixate most in regions of the picture which are considered most interesting or attention drawing. Mackworth and Morandi (1967) found that there is a good deal of similarity for fixation locations across subjects; subjects fixated upon areas considered “informative.” Considerable debate exists on what is to be considered “informative”, however, as some researchers classify objects which are “out of place” in the picture to be informative (Loftus & Mackworth, 1978; Henderson, 1992) while others consider objects that typify a scene as being informative (Antes, 1972; Mandler & Ritchey, 1977). Loftus (1983) suggests that the length of a fixation is determined by the difficulty of identifying and encoding the features fixated upon, such that easy to encode features are fixated for a shorter period than are more difficult to encode features. For instance, Loftus (1992) found that fixation durations increase for pictures whose quality has been degraded.

The above research has focused just on the eye fixations. Other research has focused on the pattern of eye fixations and saccades, also known as the scanpath. Noton and Stark (1971) found that subjects’ scanpath for the first few fixations during initial viewing is very similar to that during the second viewing. In contrast, Antes (1972) found that scanpath characteristics change across multiple viewings. At initial viewing, informative regions are fixated first and for short durations; as subjects view the picture more, however, fixation durations increased and the number of fixations on “informative” regions decreased. In addition, with greater exposure, saccade lengths decreased. Zangemeister, Sherman, and Stark (1995) also found that prior knowledge affects scanpaths; they found that the scanpaths for professional art viewers were characterized by more global, as opposed to local, viewing strategies than was found in the scanpaths of nonprofessional art viewers. Such results demonstrate that prior exposure to a picture effects subsequent processing of that picture.

Rizzo, Hurtig, and Damasio (1987) took these ideas one step further and demonstrated that analysis of the scanpaths of prosopagnosic patients could be used to determine which faces the patient had previously seen. These results were not replicated with normal subjects, however. Cohen, Althoff, Webb, McConkie, Holden, and Noll (submitted) reported that the scanpaths of normal subjects can also be used to determine which faces and complex scenes have been previously seen. Such results were obtained whether the familiar stimuli used were pre-experimentally familiar to the subject or if the stimuli were made familiar to the subject by one prior 5-second exposure. Furthermore, additional unpublished research by Whitlow, Althoff, and Cohen (1995) found that eye movement monitoring could be used to covertly detect pre-experimental exposure to a face (using faces of famous people, e.g. “Ronald Reagan”) in an amnesic subject who showed no explicit recognition of such faces. Such a result (preserved memory performance in an implicit task with poor performance in an explicit task) is consistent with previous research on memory in amnesic subjects.

Cohen, et al., (submitted) concluded that eye movement monitoring may be an effective implicit (indirect) measure of memory. For this reason the present research uses eye movements as an indirect measure of memory. Though we are referring to eye movement monitoring as an indirect measure of memory, it should be reemphasized that we (nor Cohen, et al.,) are not making claims about eye movements as an indirect or implicit measure of memory just as we are not making claims about the verbal report as an explicit measure; neither measure is process-pure for both may have “implicit” and “explicit” components to them. Each measure supplements the other for the purpose of assessing the nature of scene representations.

Overview

While previous research has examined the perceptual specificity of object representations, no research has examined the perceptual specificity of scene representations using color photos of naturalistic scenes as stimuli. We use naturalistic scenes as our stimuli because we are interested in moving beyond the issue of object representation to the issue of complex scene representation, for which comparatively little research has been conducted. In addition, by using complex scenes as stimuli, we relate the field of study on internal representations to the world of everyday experiences, such as those referred to in the opening paragraph. If we truly want to understand the degree to which an individual can recognize a place from a new view of that place, a more realistic stimulus should presumably provide greater ecological validity. By using photographs of scenes, we are still limited by the use of a two-dimensional surface that is static, but we are including richer information than is provided by line drawings or backgroundless objects.

The independent variables used in each of the five studies reported in this thesis include the within-subject variables of number of exposures (all experiments) and stimulus changes (rotation, Experiments 1-2; horizontal frame translation, Experiment 3; left-right orientation, Experiment 4), and the between subject variables of background (Experiments 1 and 2), and type of task (Experiment 3). The purpose of each of these variables is discussed in later sections.

The dependent measures in all of the present studies include the subjects' verbal reports as well as their eye movement patterns. For the verbal report, subjects verbally express their memory for the presented stimuli; since we are assessing their memory for a stimulus by asking them directly, we refer to the verbal report as a "direct measure" for the remainder of this thesis. We analyze eye movement patterns to assess the subject's prior exposure to a scene; since this information is acquired without directly asking the subject to provide any information, we refer to the eye movement measure as an "indirect measure" for the remainder of this thesis. We use eye movements because they provide a richer set of information than is provided by a simple verbal response; eye movements allow us to examine both where a subject fixated as well as how they moved their eyes. We are interested primarily in how both measures advance our understanding of scene representations, and we are less interested in how these measures differ from one another in their response characteristics, although we do address this issue in Appendix D.

In Experiments 1 and 2, we use familiarity ratings as the direct measure for three reasons. First, we wanted to use a direct measure that would allow us to note a graded change in subjects' subjective sense of familiarity with a scene over time. While a "yes/no" recognition judgement could serve as a graded measure, it would not assess subjects' subjective familiarity with each picture. Second, we wanted to investigate whether a sense of familiarity with a stimulus was related to the eye movement patterns generated towards that stimulus. Again, a "yes/no" recognition paradigm would not provide us with such information. Third, a "yes/no" recognition paradigm would have required us to equate the number of old and new items, which would have required the collection of more scenes than was practical.

In Experiments 3-4 we use “confidence of recognition” judgements, rather than familiarity judgements, as the direct measure for three reasons. First, in the psychological literature, a recognition judgement is more commonly used as a direct measure of memory than a familiarity judgement is, so using a recognition judgement allows for greater comparison with previous research. Second, a recognition judgement was used because the designs of Experiments 3 and 4 allowed for it. The designs contained fewer blocks of trials and therefore did not require a prohibitive number of scenes to balance the number of old and new scenes within each block as Experiments 1 and 2 did. In Experiments 1 and 2, the prohibitive number of scenes required to balance the design using recognition judgements forced us to use familiarity judgements, which would not require a balance between old and new scenes. Third, we used confidence ratings rather than “yes/no” recognition judgements in order to maintain a graded measure for each scene; we did not use reaction time as a dependent measure since this may have rushed the subjects and possibly added another factor affecting their eye movements.

We use eye movements as our indirect measure in all 5 experiments because it can reveal a subject’s prior exposure to a stimulus without the experimenter asking the subject to make any explicit recognition judgements. We also use eye movements because they provide much richer information than a simple verbal report can. For instance, with eye movements we can assess how prior exposure changes where the eyes move, how long they stay fixated in a particular location or object, and the pattern of movements across the picture.

General Methods

Apparatus

For the first 4 blocks of Experiment 1, images were displayed and familiarity ratings were recorded by a Macintosh Quadra with a 15” color monitor with a resolution of 756 x 480 pixels. For the last block of Experiment 1 and all other experiments reported in this thesis, a Gateway 2000 486/66 computer was used to display images to an SVGA 19” color monitor with a resolution of 756 pixels horizontally and 480 pixels vertically. Subjects sat approximately 2 feet from the monitor. Eye movement data were collected using an ASL Series 4000 Eye Tracking System. Fit to each subject’s head was the optics device that monitored the position of the left eye, an infrared illuminator, and the magnetic field sensor that determined head position by recording the strength of the magnetic field being generated by a magnet fixed directly behind the subject.

The position of the eye was determined by noting the distance between the first Purkinje image (corneal reflection) created by the infrared illuminator and the center of the pupil. Eye position was measured at a rate of 60Hz. The accuracy of the eye tracker is appoximately one-half a degree of visual angle.

Eye Movement Variables and Data Analysis

Analysis of the eye movement data proceeded as follows: using a program provided by ASL, raw eye movement data was transformed into fixation positions and durations. This fixation information was then analyzed by EMTOOL, an eye movement analysis program developed in Neal Cohen's laboratory (Maciukenas, Althoff, Holden, Webb, & Cohen (1997), which determined for each picture trial the values of each of the eye movement variables listed below. All of the eye movement variables used are used because Cohen, et al., have found these variables to be effective indicators of prior exposure.

The eye movement variables examined in these experiments included fixation variables, which concern the duration and/or location of fixations, and transition variables, which encode the structure or degree of constraint present in a scanpath. The fixation variables include the following:

> (nfix)--the number of fixations per trial. This variable has been used by numerous researchers (Loftus, 1983; Antes, 1972). Since early processing involves derivation of the overall "gist" of the scene rather than analysis of scene details (Antes, 1972), more fixations are generated when familiarity with the scene is low.
>
> (mfd)--the median fixation duration per trial. This variable has been used in numerous eye movement studies. Loftus (1983) found that mfd was effected by encoding difficulty while Antes (1972) found that mfd was effected by the amount of exposure to the picture. As more detailed analysis occurs later in processing of a scene, median fixation duration becomes longer with greater familiarity.
>
> (nclust)--the number of fixation "clusters." A cluster is a region of space of unlimited size that is composed of fixations that are within 30 pixels of each other. If no other fixations are within 30 pixels, then the cluster consists of only one fixation. Clusters represent the number of places a person looks, rather than the number of fixations.

Such a measure may be a better indicator than nfix of whether a subject is attempting to derive the overall “gist” of the scene (resulting in more clusters) or doing a more detailed analysis (resulting in fewer clusters) since one cluster may be composed of numerous fixations.

(leftfix)--the proportion of fixations that fell on the left side of the picture. This variable is used because research by Cohen, et al., (submitted) and Althoff and Cohen (in manuscript) has found this variable to be an effective indicator of face familiarity. Face recognition relies more strongly on the right hemisphere than the left (Rapcsak, Polster, Gilsky, & Comer, 1996; Sergent & Signoret, 1992) . This assymmetry may be the basis for the effectiveness of this variable, and to the extent that scene recognition also relies on a symmetrical processing, this variable may prove effective for scenes as well.

(lefttime)--the proportion of time spent on the left side of the picture. This variable is similar to leftfix.

(retfix)--the number of fixations that occurred before a fixation returned to the original fixation location. This variable is used because the eyes may return to a starting location at an earlier point in the scanpath of an unfamiliar object than a familiar object if it is the case that the derivation of the spatial relationships between objects is a more dominant purpose for unfamiliar objects than familiar ones. In addition, we expect that the number of fixations vary for unfamiliar versus familiar scenes.

(rettime)--the amount of time that passed before a fixation returned to the original fixation location. This variable is used for the same reasons we use retfix and because we expect familiarity to affect the median fixation duration.

The transition variables include 8 entropy measures that represent the degree of constraint or nonrandomness present in the eye movement path. These values are derived by constructing transition tables that represent the eye movement path and determining the "informativeness" of this table. A table is constructed by specifying with one's the number of times the eye moves from one cluster (rows) to another cluster (columns). If an eye movement path were completely random and unconstrained, the table would be uninformative, as one's chance of predicting the location of a subsequent fixation given the location of the present fixation would be low. Conversely, an "informative" table corresponds to a higher level of constraint (higher values in some columns than others), or predictability. The 8 entropy measures are described below. All h variables were used previously by Cohen, et al., (submitted), Althoff and Cohen (submitted), and Whitlow, Althoff, and Cohen (1995).

(h1)--a first-order entropy measure, which is a measure of the likelihood of predicting the location of a subsequent fixation given the location of the present fixation. Rizzo, et al (1987), used a measure similar to this one and found that it distinguished between the faces prosopagnosics had and had not seen before.

(h1t)--a first-order entropy measure weighted by the duration of the fixations at each location.

(s1)--a first-order entropy measure, normalized for the total number of cells in the

Markov matrix that describes the scanpath for a given trial. This and other s variables were added to remove the influence of number of fixations which is present in the h entropy measures, so each s measure is a normalized version of its h measure counterpart.

(s1t)--a first-order entropy measure, weighted by the duration of the fixations and normalized for the number of fixations in the path.

(h2, h2t, s2, s2t)--identical to those four described above except that they measure the likelihood of predicting the location of two subsequent fixations given the location of the present fixation.

In previous research (Cohen, et al., submitted; Althoff and Cohen, in manuscript; Whitlow, Althoff, and Cohen, 1995) familiar images display higher entropy values, lower median fixation durations, more clusters, and a greater number of fixations than a path that is unfamiliar, or constrained. The statistical similarity among these variables, as well as a principal components analysis of all of them based on data from Experiment 2 can be found in Appendix A.

Since there are so many eye movement variables, an attempt to simplify the presentation of these measures was made by constructing 2 sets of tables for most ANOVAs. The first table will normally provide the results from the direct measure and some of the more robust eye movement variables (nfix, h2t, nclust), while the second table of the set will normally provide the results from the direct measure and all of the eye movement variables. The variables within this smaller table are heretofore referred to as the “primary variables.”

Hypotheses

The primary aim of these experiments is to investigate the nature of scene representations by examining their perceptual specificity for various types of changes. We hypothesize that scene representations are perceptually specific for view and reflection, but not translation. We make the above hypotheses because while previous research (discussed above) has found scene representations of objects to be perceptually specific for view and reflection, no research has demonstrated sensitivity for translation. Research by Intraub, et al, (1996) in which no perceptual specificity was found in relation to picture frame changes, is the basis for our predictions for Experiment 3, in which scenes are translated. Therefore, for Experiments 1, 2 and 4 we expect to disprove the null hypothesis and find a significant difference in both measures between the values for the unchanged stimulus and the changed stimulus. For Experiment 3, we expect to not disprove the null hypothesis and hence find no significant difference between unchanged and changed stimuli.

More specific hypotheses relevant to each particular experiment will be discussed after the Methods section for each experiment.

Experiment 1: Multiple Views of Naturalistic Scenes

Design

A within-subjects, two factor (block x condition) design was used for this experiment. Familiarity for the scenes was increased across time (five blocks) by either repeating the same view, and hence the same image of a particular scene (same condition), or by displaying a different view of that specific scene (different condition).

In order to better test the nature of the representations constructed across exposures, the same and different conditions were each divided into two sub-conditions; thus, there were four experimental conditions and one control condition. In each of the five blocks 25 images were shown, 5 from each of the 5 conditions. In the same-same condition, subjects saw the same view of the scene in all 5 blocks. In the same-different condition subjects saw the same view of the scene for the first 4 blocks and saw a new view of the scene in the 5th block; the new view selected was counter-balanced for each subject such that each of the 5 views were represented an equal number of times. In the different-different condition, subjects saw a different view of the scene in all 5 blocks, whereas in the different-same condition, subjects saw a different view of the scene in the first 4 blocks and saw in the 5th block the same view shown in block 1 for that scene. In the control condition, subjects saw completely new scenes in all 5 blocks. Depending on the type of analysis being performed, the first 4 blocks are treated as having either 3 conditions (same, different, control), or 5 conditions (same-same, same-different, different-different, different-same, control).

We use five blocks of exposure because we want to make these unfamiliar scenes as familiar as a famous scene, such as the Golden Gate Bridge or a subject's own back yard. As previous research by Cohen, et al., (submitted) had compared famous (pre-experimentally familiar) stimuli (e.g. Ronald Reagan) with unfamiliar stimuli, we want to make our pre-experimentally unfamiliar stimuli comparable in familiarity to the famous stimuli in the Cohen, et al., research. In addition, by using 5 blocks of trials we wished to determine how much exposure is necessary for our familiarity ratings and eye movement measures to reach a plateau, or "maximum familiarity" in terms of these measures.

Method

Stimuli. The stimuli used in this experiment included 45 naturalistic scenes from around rural Illinois and Indiana. The scenes were videotaped using a VHS camcorder and later digitized into 640 x 480 pixel Pict files using the AVID program on a Macintosh computer. Using the Debabbelizer program for the Macintosh, the images were then converted into 16-bit indexed color Pict files for use on the Macintosh Quadra. All images were also converted into Targa file format for use on the Gateway 2000.

For each scene, five images from different views were taken (Figure 1.1); for each of the images, an effort was made to ensure that the center focus for each image was the same both vertically and horizontally, that the distance from the videocamera to the scene's center was the same, and that the difference in angle between each image was 15 degrees. Relatively accurate fifteen-degree rotations were obtained by taking a camera and a degree-marked tripod to the horizontal center of the scene in the to-be-captured pictures and zooming in on and marking the locations from where the to-be-captured pictures would be taken. An effort was also made to ensure that no single view contained an object in the foreground that occluded the "scene" or that made that particular image of the scene significantly different than the others. Each scene contained a mixture of natural and human-made objects such as houses, school buildings, churches, pavilions, bridges, sheds, swingsets and so forth (Figure 1.2).

The full stimulus set for each experiment included 225 pictures. These 225 pictures were made up of 45 different scenes with 5 pictures taken from different views of each of the 45 scenes. The displaying of each scene was completely counterbalanced such that each subject saw all 45 scenes, and that across subjects, all scenes, as well as all views of each scene, appeared in all conditions an equal number of times. The view images selected across scenes were fully counterbalanced within each subject. Each instance of a scene showed up only once per block.

Participants. Forty-five University of Illinois students (16 males, 29 females), ages 18 to 28, were paid 5 dollars for participating in this study. All subjects reported normal or corrected-to-normal vision.

Procedure. Subjects were instructed to examine each scene and make a familiarity judgment ("How familiar does the scene feel to you?") about the scene on a 1-7 scale (with "1" as unfamiliar). Subjects were also told, "Over the course of the experiment, you will see scenes from various views. Some scenes shown in one view during one part of the experiment can reappear in the same or different view in a later part of the experiment. Some scenes are shown only once." During the subsequent practice session, which included 10 images that were only displayed during this practice session, the experimenter gave feedback to the subject to ensure that the subject understood that familiarity judgements were to be based on the general scene and not the exact picture. If subjects recognized the scene and also realized that the viewpoint was different, the subject was told to ignore the change in viewpoint and rate the scene as being highly familiar. After the practice session, the experiment began. Each trial began with a fixation point in the center of the screen for 1.5 seconds, followed by the stimulus, which remained for 5 seconds, followed by a four-second prompt to make a familiarity judgment. The subject recorded his/her familiarity judgments by pressing a number (1-7) on the keyboard. In order to eliminate primacy and recency effects, each block contained two scene images prior to and after the 25 images described earlier; these same scenes served only as primacy and recency buffers for all subjects, and although data was collected while they were displayed, this data was not analyzed. Subjects received a short break after each block.

After the 4th block, subjects were taken to another room where eye movements were recorded. Once the eye and head position were calibrated, the trials continued as before, except that the subjects would verbally express their familiarity judgments while the experimenter wrote them down. During this block, the fixation dot would not be replaced by the image until the experimenter confirmed with a button press that the subject's eyes were fixated on the dot.

Hypotheses

As was stated in the General Methods/Hypotheses section, we believe that scene representations are perceptually specific to rotations in depth and therefore expect to falsify the null hypothesis by finding a significant difference in both of our measures between images which remain the same (same condition) and those which change (different condition). The finer details of this hypothesis for both of the measures are outlined below.

Direct measure. In this experiment, conditions in which viewpoint changes across exposures should show significantly different effects than those conditions in which viewpoint does not change. Thus, we expect that performance on the direct measure will be greatest when viewpoint remains the same (same condition). We expect this sort of result because previous research has found such effects with objects and because we believe that performance on this measure is mediated by view-dependent representations. If this is so, the perceptually specific effects should be a function of rotation from the nearest learned image(s).

If, instead, more abstract representations are constructed, then we would expect the following four perceptually nonspecific results. First, no significant difference would emerge between the familiarity ratings for the same condition and the different condition, since viewpoint should not matter, though a trend towards lower familiarity may arise due to occlusion of object parts. It should be noted that a result of increasing familiarity over time in the different condition does not imply an abstract representation, as viewer-centered representations may be tuned to a range of views that may overlap. Second, familiarity ratings for images in the same condition when a new view is shown in block 5 should not change significantly. Third, a decrease in familiarity for the same-different images should not be directly related to the amount of rotation between training and test views. Fourth, in the fifth block, familiarity ratings for images in the different-same condition should not be significantly different than familiarity ratings for images in the different-different condition, as showing the same view again should not be facilitative.

Since we expect that the representations are primarily view dependent, however, we expect to disprove the above hypotheses and find perceptually specific effects. Thus, we first expect to find a significant difference between familiarity ratings for the same and different conditions. Second, we expect to find a significant decrease in familiarity when images trained in the same condition change views and that the resultant familiarity for scenes in the same-different condition in block 5 should be significantly less than the resultant familiarity for new views of scenes shown previously in different views (different-different condition). This should be the case since the same-different condition averages 30 degrees of rotation while the different-different condition changes by only 15 degrees. In addition, we expect that this decrease in familiarity in the same-different condition should be a function of rotation from the training view. The reasoning behind this second set of predictions follows. Since the views of each scene are separated by 15 degrees, as long as the view-dependent representation is broadly tuned to something greater than 15 degrees in each direction from training view, a new view presented in block 5 in the different condition will have more “coverage” than a new view presented in block 5 in the same condition. So even though we believe that view-dependent representations underlie both, we expect to see more perceptually specific effects to emerge from the same condition than from the different condition. Lastly, we predict that familiarity for scenes in the different-same condition will benefit from the re-use of the representation developed in block 1 and rise above the familiarity levels of the different-same condition in block 5.

Indirect measure. Assuming that scene representations are perceptually specific for rotations in depth, we expect to find a similar pattern of perceptually specific effects with our eye movement measures than we did with the direct measure. Thus we hypothesize that in conditions in which the stimulus does not change, eye movement variables such as nfix, nclust, and entropy measures should show significantly lower (i.e. more “familiar”) values than those in which the stimuli do change. The specific pattern of results we expect is identical to that which was outlined for the direct measures above.

Results

Because the first four blocks consisted of basically three conditions (same, different, and control) while the 5th block consisted of five conditions (same-same, same-different, different-different, different-same, and control), familiarity ratings in the first four blocks were analyzed separately from familiarity ratings in the 5th block. In order to equate for number of items within each condition for the first four blocks, the mean value for the same condition was determined for each subject by averaging across 10 trials (5 same-same and 5 same-different), as they were conceptually identical during these blocks. The same procedure was performed for the different condition trials.

In the first four blocks, a 4 (Block) x 3 (Condition: same, different, control) repeated measures ANOVA was performed across the 45 subject means. For block 5, a 2 (Block) x 5 (Condition) repeated measures ANOVA was performed across the 45 subject means.

Planned comparisons were performed between specific conditions.

Same view versus different view for the direct measure. The present experiment examined the perceptual specificity of scene representations as measured by subjects' familiarity ratings for rotation of scenes across exposures. We first hypothesized that familiarity for scenes exposed repeatedly in the same view (same condition) would increase more rapidly than familiarity for scenes exposed repeatedly in a different view (different condition).

For this part of the experiment, data in the same-same and same-different conditions were collapsed together to make the same condition; the different-different and different-same conditions were collapsed into one different condition. As can be seen in Figure 1.3, the familiarity for scenes in the same condition rose faster than the familiarity for scenes in the different condition.

A repeated measures ANOVA was performed and found a significant main effect for condition, block, and condition by block interaction (see Table 1.1). To determine if the different condition ever reaches the same level of familiarity as the same condition, a planned comparison was performed between the ratings of the different and same conditions in block 4, yielding a significant difference [$F(1, 44)=4.9$; $p<.03$]. Hence, scenes shown 4 times from different views did not quite reach the level of familiarity of scenes shown 4 times from the same view.

These results support our first hypothesis that images viewed in the same view at each exposure will show a steeper increase in familiarity than images viewed in a different view at each exposure. Such perceptually specific results suggest that one’s sense of familiarity is mediated by a representation with view dependent components.

Familiarity for a new view in the same and different conditions. My second major hypothesis stated that familiarity for scenes shown in a new view in the fifth block will be significantly less if previous exposure to that scene had been from the same view (same-different condition) than if previous exposure had been from different views (different-different condition).

The rightmost part of Figure 1.3 shows the results for the fifth block. As one can see, the same-same condition shows approximately the same overall rating (6.5) as the different-different (6.3) and different-same (6.5) conditions, while the same-different (5.7) condition falls almost a point from its rating in block four.

The data for block 4 were divided into the five conditions that are present in block 5, and a repeated measures ANOVA was performed on these 5 conditions in these two blocks. A significant main effect of condition, block, and condition by block interaction was found (see Table 1.1). Individual contrasts performed on conditions across blocks 4 and 5 revealed a significant drop in familiarity for the same-different condition [$F(1, 44)=64.1$, $p<.0001$], consistent with our original hypothesis; no other conditions showed a significant change across blocks. Individual contrasts between conditions in block 5 reveal a significant difference ($p<.05$) between the same-different condition and all other conditions, which is also consistent with our prior predictions. Thus, the primary finding that emerged from these analyses on blocks 4 and 5 is that the familiarity for scenes in the same-different condition falls significantly below the other experimental conditions. Thus, our second hypothesis was supported; performance in the same-different condition was more perceptually specific than performance in the different-different condition. Since the same-different condition represents an average rotation of 30 degrees while the different-different condition represents a 15 degree rotation from the nearest learned view, the lower familiarity in the same-different condition represents a view-specific effect.

To investigate whether the decrease in familiarity in the same-different condition in block 5 is more consistent with what one would expect from a view-dependent versus a view invariant representation, mean familiarity ratings were obtained for scenes which changed by 15 degrees, 30 degrees, and 45 degrees. An inverse relationship between the amount of view change and familiarity would be supportive of a hypothesis that stated that scene representations have view dependent components. To test this hypothesis, a repeated measures ANOVA was performed using the 45 subject means of these three rotation conditions and the same-same and different-different conditions. Figure 1.4 provides a graphic illustration of the effects of rotation. As Table 1.2 shows, significant effects of rotating the scene begin to show up between 15 and 30 degrees of rotation from the nearest learned viewpoint. This is a conservative analysis since the means were composed of all trials, including those in which rotation did not have an effect (ratings of “7”). A more robust effect of rotation might have emerged if we had not included those trials with “7” ratings (52% of the trials), but this could not be done due to an excessive number of empty cells. These results demonstrate that while representations have a breadth of coverage of at least 45 degrees, the significant differences in familiarity that emerge between 15 and 30 degrees support the argument that view-dependent representations are mediating a subject's sense of familiarity. Furthermore, the broadly-tuned nature of these representations is noted by the result that the different-different condition (which represents a 15 degree rotation) is significantly different than the 30 degree rotation in the same-different condition, while no other 15 degree differences emerge as being significant. Presumably this is so because the different-different condition forms multiple representations that in many cases "surround" the new view while the new view in the same-different condition is adjacent to only one representation.

In summary, the familiarity ratings in block 5 show perceptually specific responses, which is supportive of a hypothesis that scene representations are view dependent.

<u>Analysis of variance of the indirect measure.</u> To determine what the eye movement measures reveal about the nature of the scene representations, a repeated measures ANOVA was performed on the 45 subject means of the 5 conditions in block 5. This was performed for every eye movement variable. Table 1.3 presents the ANOVA table for all of the variables. Table 1.4 lists the F values and Mean Squares for all of the variables for the between-condition comparisons that proved significant. These tables, as well as Figures 1.5 to 1.18, clearly show that for almost all of the variables, a significant difference exists between familiar and unfamiliar scenes. In addition, the pattern of the results for each variable follows the same trend that is seen in the familiarity data, demonstrating a very high degree of sensitivity of the eye movement measures. That each of the eye movement variables can distinguish between conditions whose mean familiarity ratings were within a

point of each other is certainly noteworthy.

Regression analyses were performed to evaluate the relationship between familiarity ratings and eye movement variables. These analyses, which are reported and discussed in Appendix B, revealed strong relationships between the two sets of measures. To determine whether eye movement variables alone could be used to detect whether a subject has been exposed to a scene or not, discriminant analyses were performed, and these results are reported and discussed in Appendix C.

Retention data. To examine the effect of training (same view versus different view) on memory retention, we had 17 subjects return 7 months later. Each subject was shown 5 scenes from the same-same condition (in the same view again) and 5 scenes from the different-same condition (in a different view they had not seen). Subjects were asked to rate the familiarity of these scenes on a 1 to 7 scale, with 7 as most familiar. The same-view and different-view scenes had identical means (5.4) and standard deviations (1.2) and were significantly different than the control scenes.

These results clearly demonstrate that subjects can maintain, over a 7 month period, a high level of familiarity for scenes for which they had a great deal of exposure to initially, even if the scene is shown in a new view. These results also suggest that perceptually specificity for view declines over time, due to the converstion of view-dependent representations being converted into more abstract representations (Lawson & Humphreys, 1996). This result is consistent with what one should expect in terms of what is most adaptive for the organism; since the specific details of scenes change over time and since one is unlikely to view a scene from the same vantage point twice, it is advantageous if the representations generalize over time.

Discussion

The familiarity ratings from Experiment 1 yield two important results that both support the argument that scene representations are primarily view-dependent. First, familiarity for color images of complex scenes increases at a faster rate when the same view of the scene is repeatedly shown than when different views of that scene are repeatedly shown. Second, the familiarity ratings for the scenes in the same-different condition in block 5 fall significantly below the familiarity ratings in the other experimental conditions, including the different-different condition. In particular, trial by trial analysis of familiarity ratings indicates that familiarity ratings for scenes in this condition in block 5 varied inversely with the amount of view-rotation between it and the view displayed in the previous 4 blocks.

This pattern of results is consistent with the view that scene representations are primarily view-dependent. The graded decrement in familiarity as a function of view rotation is consistent with electrophysiological research and computational models of object recognition across views (Vetter, et al., 1995).

The eye movement variables also provided support to the argument that scene representations are view-dependent for their pattern of results (from the ANOVA analyses) resembled those of the direct measure. Eye movement values in block 5 paralleled familiarity ratings even though familiarity ratings differed seemingly little from each other between conditions. This similiar pattern of results was supported by the regression analyses reported in Appendix B, in which more than two-thirds of the variables showed a significant ($p<.05$) r square value, including mfd with an r square value of .93 (Table B.1).

The familiarity ratings test performed 7 months after initial exposure demonstrated that subjects can maintain memory for scenes over a long time period, at least for scenes with which they had already attained a high level of familiarity. The identical ratings of the same and different conditions suggests that there is no effect of the type of training on retention and that the representations developed for the scenes in the different condition were either view invariant or view-dependent with wide "coverage." It is impossible to say which of these representations mediated performance in the different condition; we expect that had we shown a new view of scenes in the same condition, however, the view-dependent representations that were formed in same condition would have generated lower familiarity ratings, just as they did in the same-different condition in block 5. These results may also have been different if these two conditions had not already achieved identical levels of familiarity during the initial set of exposures.

This is one of the first experiments to have investigated the view-specific nature of the representations of complex, naturalistic, outdoor scenes using different objects in each scene; Diwadkar and McNamara's (1997) study was similar but it used indoor objects and each object was used in each scene. It is also the first experiment to have used eye movements to study this issue. Our results are similar to those results obtained in research investigating the issue of view-specificity of representations of objects and so we conclude that scene representations are view dependent.

Though we have shown with both variables that view-dependent representations are mediating our results, we still know little about the nature of scene representations. Were these results unique to complex scenes or would they have obtained with simple objects? Are the scene representations that we refer to simply the summation of the representations of all the subordinate objects within the scene, or is the scene represented as a unique complex superordinate object in and of itself? Would we obtain the same results if we had only rotated the foreground objects or the background objects relative to the observer? In Experiment 2, we attempt to address these questions by removing the background in order to examine the effects of rotating only a foreground object.

Experiment 2: Multiple Views of Naturalistic Scenes with Backgrounds Removed

Method

Stimuli. The stimuli used in this experiment included the same 45 naturalistic scenes used in Experiment 1, except that their backgrounds were removed and replaced with black, using Adobe Photoshop. Backgrounds were removed in such a way so as to leave intact the object in each scene that had been chosen during the original videotaping of the scenes as the structure of focus in the picture (house, church, pavilion, etc.). For pictures in which an object such as a tree or a bush occluded the object of focus, the image was edited in such a way so as to make it appear that the occluding object was never present in the first place. Hence, the only differences between the five views of each scene are those differences that arise from the rotation of the focus object. All images were converted into Targa file format for use on the Gateway 2000.

Participants. Forty-five University of Illinois students (16 males, 29 females), ages 18 to 29, were paid 5 dollars for participating in this study. All subjects reported normal or corrected-to-normal vision.

Procedure. Given the usefulness of the eye movement measures in Experiment 1, we decided to monitor eye movements during all of Experiment 2 to see how they change across time. Since eye movements were recorded during all 5 blocks in this experiment, the procedure for this experiment is almost identical to that of the 5th block of Experiment 1. The only differences between the first four blocks of Experiment 1 and the present experiment are that the subject’s eye and head position were calibrated immediately after the practice session, the fixation dot was replaced by the image once the experimenter confirmed with a button press that the subject’s eyes were fixated on the dot, and the subject would verbally express his/her familiarity judgments while the experimenter manually recorded them.

Hypotheses

Perceptual specificity. Hypotheses regarding the perceptual specificity of the representations and their consequent effects on the direct and indirect measures in this study are identical to that of Experiment 1. Since both prior research and Experiment 1 have found representations to be perceptually specific for viewpoint, we expect the same result to emerge here with the backgrounds removed from the scenes. See Experiment 1 for more specific details of what a perceptually specific effect entails.

Background. Since the only difference between Experiments 1 and 2 is the presence or absence of background, this study examines the role of background in scene recognition. If scene recognition was based primarily on the recognition of the primary object within a scene, then the presence or absence of background should have no effect on familiarity ratings or eye movement variables. In other words, to the extent that scene recognition is based on the recognition of the primary object in the scene, there should be few differences in familiarity ratings and eye movement measures between these two experiments; to the extent that the background plays a role in scene recognition, the differences should be enhanced. We expect that it does play a role. Specifically, the presence of a background requires more complex encoding, including that of multiple objects and the spatial relationships between them; thus, familiarity in the same condition should increase faster when no background is present. In the different view condition, where both the object and the scene changes with the background present and only the object changes with the background absent, the difference between the increase in familiarity between experiments should be even greater.

For the eye movement measures, the presence of a background should play an important role. Since the normal pattern of picture viewing involves a quick scanning of the primary features/objects in a scene, followed by a more detailed examination of sites of interest (Antes, 1972), the absence of a background should reduce or eliminate the initial stage of viewing. To the extent that the eye patterns for “familiar” scenes also involve less processing in the initial scanning stage, the removal of a background should have the effect of shifting both familiar and unfamiliar scenes in the “familiar” direction. If a limit exists in how “familiar” scanpaths can become, then such a shift should have the effect of making discrimination between familiar and unfamiliar scenes more difficult. In addition, as there are fewer objects to fixate on and less spatial processing to be done when backgrounds are absent, it is quite likely that eye movement variables optimal for predicting familiarity for background-present scenes are not the same variables that are optimal for background-absent scenes.

Results

The primary purpose of Experiment 2, like Experiment 1, was to determine the perceptual specificity of the representations for rotation of outdoor objects. We hypothesized that perceptually specific effects will be obtained for both types of measures. For this analysis, we divide the 5 blocks of data into two parts, as we did in Experiment 1. The first part consists of the first 4 blocks where we compare results across the same, different, and control conditions to determine whether repeated exposure of a scene from different views affects our dependent measures differently than repeated exposure of the scene from the same view. The second part consists of blocks 4 and 5 where we compare results across the same-same, same-different, different-different, different-same, and control conditions, to determine how the interaction of new view and training view effect the two types of measures. The mean for each condition for each subject is derived from 5 trials of pictures, except for the same and different conditions in the first four blocks which are derived from 10 trials of pictures (just as in Experiment 1). Unlike Experiment 1, familiarity and eye movement data were collected for both parts, not just block 5, and hence will be reported and discussed together.

The purpose of Experiment 2 is to examine whether the perceptually specific characteristics of scene representations (examined in Experiment 1) are similar to the characteristics of object representations and to determine whether the effects observed in Experiment 1 with complex scenes are mediated primarily by the primary foreground object rather than background objects or the scene as a whole.

To determine if the removal of background in the scenes alters the pattern of results obtained in Experiment 1, we treat background as a between subjects variable and compare the direct and indirect measures between experiments; specifically we compare the direct measures for the first four blocks and the 5th block alone and the indirect measures for the 5th block alone (since eye movement data was not collected in the first 4 blocks of Experiment 1).

First four blocks. Since we hypothesized that perceptually specific effects would be obtained, we expected that in the first 4 blocks, scenes exposed repeatedly in the same view (same condition) would show a different pattern of results than scenes exposed repeatedly in a different view (different condition). Specifically, we expected that familiarity ratings would rise more rapidly in the same condition and eye movement variables would change more rapidly in the same condition.

Figures 2.1 through 2.4 illustrate the results for the familiarity ratings, nfix, h2t, and nclust, respectively, which all demonstrate the expected results stated above. Figures 2.5 through 2.10 illustrate the results on other indirect measures that show weaker or a lack of effects. Repeated measures ANOVAs were performed on both the direct and indirect measures. A significant main effect for condition, and condition by block interaction was obtained for the direct measure and a significant main effect for condition, block, and condition by block interaction was obtained for most of the indirect measures. Table 2.1 provides the ANOVA results for 4 selected variables while Table 2.2 provides the ANOVA results for all variables. Individual contrasts were performed between the three conditions (same, different, control) in each of the 4 blocks and the ANOVA results for this analysis for all variables is in Table 2.3. Comparison of the figures and tables demonstrates that the variables with the highest F values (nfix, h2t, the “fixation” factor, and nclust) are also the variables whose pattern most closely resembles the pattern of the familiarity ratings. These results, in which the direct and indirect measures change more rapidly in the same condition than in the different condition, support the primary hypothesis concerning perceptual specificity for view. Such a perceptually specific result is supportive of the argument that scene representations are view dependent.

To determine whether familiarity for scenes in the different condition reaches the level of familiarity for scenes in the same condition, a planned comparison was performed between the ratings (direct measure) of the different and same conditions in block 4, yielding a significant difference [$F(1, 44)=24.7$; $p<.0001$]; hence, after four blocks, scenes in the different condition do not attain the familiarity level of the scenes in the same condition.

<u>Blocks 4 and 5.</u> In the final block of this experiment we examined whether the type of view exposure (same or different) in the previous four blocks interacts with the type of view (same or different) shown in block 5. In particular we are interested in whether showing a new view of a scene seen only in one view has the same effect on our dependent measures as showing a new view of a scene seen in 4 different views. Since the only difference between this experiment and Experiment 1 is the removal of background, we expect to replicate the perceptually specific, view-dependent results that we found in Experiment 1. Specifically we expect that values in both measures in the same-different condition will show a greater change between blocks than they do in the different-different condition since the former averages a 30 degree rotation and the latter represents is a 15 degree rotation. For the direct measure, this change will be a decrease in familiarity while for most indirect measures it will be an

increase. In addition, we expect the different-same condition to show a greater increase in the direct measure and greater decrease in the indirect measures compared to the different-different condition.

The rightmost part of Figures 2.1 through 2.10 shows the results for the fifth block for all measures. For the direct measure (Figure 2.1) the same-same condition shows approximately the same overall rating (6.9) as the different-different (6.6) and different-same (6.7) conditions, while the same-different (6.4) condition falls almost half of a point from its rating in block four, thus replicating the expected pattern of results. This pattern of results also emerged for some of the indirect measures (nfix, h2t, “fixation” factor, nclust), though with greater variability.

To determine the significance of these results, repeated measures ANOVAs were performed on the 5 conditions in blocks 4 and 5 for both the direct and indirect measures. A significant main effect for condition and a block by condition interaction was found for the direct and some of the indirect measures (see middle section of Table 2.1 for the primary variables and Table 2.4 for all variables). For the direct measure, a planned comparison for the same-different condition between blocks 4 and 5 revealed a significant decrease in familiarity [F(1,44)=23.5, p<.0001], which is consistent with prior predictions. Other contrasts performed within conditions between blocks revealed a significant rise in familiarity for the different-different condition [F(1, 44)=11.4, p<.001] and the different-same condition [F(1, 44)=13.2, p<.001]. Thus after 4 blocks, familiarity with the scenes in the different condition is still rising.

To determine the significant differences between conditions within block 5, individual contrasts were performed between conditions within block 5 alone for both the direct and indirect measures. Tables 2.5 and 2.6 list the comparisons for selected measures, using data from blocks 4 and 5, and block 5 alone, respectively. A significant difference was found with most of the comparisons for the familiarity and nfix variables, and all the comparisons with the control condition for the other indirect measures. In particular, for the direct measure we found that the same-different condition falls below all other conditions, and that scenes shown repeatedly in different views for 5 blocks (different-different) never reach the level of familiarity as those shown in the same view for 5 blocks (same-same). In addition, we found no significant differences between the different-different condition and the different-same condition. Thus, our prediction that the same-different condition would fall significantly below the different-different condition (as it did in Experiment 1) was accurate, while our prediction that the different-same condition would rise significantly above the different-different condition (as it did in Experiment 1) was not borne out.

As Table 2.6 shows, nfix, retfix, and rettime are the only variables for which significant differences between non-control conditions were found in block 5. Nevertheless, as in the first four blocks, comparison of the ANOVA tables with the figures demonstrates that the indirect measures with the highest F values in block 5 are the ones whose pattern of results most closely resembles the pattern of the direct measure. Though many of these differences turn out to be nonsignificant due to variability, the fact that the same pattern exists for the means is noteworthy.

These results for the direct measure and, to some extent, the indirect measures are consistent with our second hypothesis that the type of exposure in the first four blocks interacts with the type of view shown in the fifth block, just as in Experiment 1; finding a more perceptually specific result in both measures for the same-different condition than the different-different condition is supportive of the argument that scene representations are view dependent. The fact that the different-different condition shows a less perceptually specific result is due to it rotating 15 degrees less than the the average same-different condition image.

Since previous research by Cohen, et al., has utilized pre-experimentally familiar (famous) stimuli to compare with unfamiliar stimuli, we were interested in determining that point in time at which the values of our direct and indirect measures no longer change with more exposures, reaching a sort of "maximum familiarity" in terms of what we can measure. Visual inspection of Figures 2.2 through 2.10 demonstrates that all of the eye movement variables in the same-same condition reach their maximum difference from the control condition in block 4, just as the familiarity ratings do. Individual contrasts performed on each variable for the same-same condition between blocks 4 and 5 support this observation by finding no significant differences for any of the variables. Thus, our familiarity ratings and our eye movement measures suggest that our scenes reach "maximum familiarity" after 4 blocks of 5-second exposures.

The effect of background. Finally, we examined the effect of background on our results. We hypothesized that due to the absence of background objects and between-object spatial relationships in Experiment 2, familiarity for the same and different conditions would increase more rapidly in Experiment 2 than it did in Experiment 1. Figure 2.14 provides a graphic comparison of the familiarity ratings from both experiments; note that for purposes of clarity, only the same-same, different-different, and control data are displayed for block 5 of this figure. In order to determine whether the presence of a background had an effect on familiarity ratings in the first four blocks, a 2 (Background) x 4 (Block) x 3 (Condition) ANOVA was performed with background (experiment) as the between-subjects variable, and block and condition as the within-subject variables. All possible main effects and interactions (including three-way) were found to be significant, with the exception of a lack of a main effect of background (see Table 2.7).

To determine whether there were any significant differences between experiments, individual t-tests were performed. These tests revealed that while the same condition in Experiment 2 shows a faster increase in familiarity than the same condition in Experiment 1, it only approaches significance in blocks 3 [$t=1.7$, $p=.10$] and 4 [$t=1.9$, $p<.07$]. The control condition in Experiment 2 maintains a consistently lower familiarity level than in Experiment 1 but is only significant in blocks 1 [$t=2.3$, $p<.02$), 3 [$t=2.0$, $p<.05$] and nearly 4 [$t=1.8$, $p=.08$]. The different condition shows no significant differences between experiments. The lack of differences between experiments is contrary to what we had hypothesized and suggests that the primary foreground object holds a greater weight in a complex scene's representation than we expected.

For block 5, a Background x Condition ANOVA was performed on all five conditions for both the direct and indirect measures. This analysis yielded a main effect of background and condition for the direct and most indirect measures, as well as a background by condition interaction for the direct measure (Table 2.8). For the direct measure in the same-same condition, background-present scenes had a significantly lower level of familiarity (6.5) than background-absent scenes (6.9); for the same-different condition, background-present scenes had a significantly lower level of familiarity (5.7) than background-absent scenes (6.4); and for the different-different condition, background-present scenes had a significantly lower level of familiarity (6.3) than background absent scenes (6.6) (see Figure 2.14). It is noteworthy that changing the view for images trained in the same view had more of an effect on familiarity for scenes with the background present. It is also noteworthy that familiarity for background-present scenes in the different-different condition seemed to level off between blocks 4 and 5 while familiarity for background-absent scenes continued to rise between blocks 4 and 5.

To summarize, examination of the interaction effects for these analyses on the direct measure reveals no significant differences between experiments until block 5, where background-absent scenes show significantly greater familiarity than background-present scenes for the same and different conditions. In the first four blocks, background had a slight effect in the same and control conditions, but not in the different condition; for the same condition this effect was in the direction of greater familiarity, while for the control condition this effect was in the direction of lesser familiarity. In addition, the effects of changing the view for same condition scenes in block 5 had a more detrimental effect on background-present scenes. Lastly, the standard error of the familiarity ratings was less with background-absent scenes. These results seem to suggest that the absence of background in a scene makes it only slightly easier to determine that an old scene is old and that a new scene is new and suggests that the representations of the scenes are not very different between experiments.

For the indirect measures, repeated measures ANOVAs were performed on the 45 subject means of each of the variables in block 5 of both experiments (see Table 2.9). These tables demonstrate that a significant difference exists between the eye movements generated towards scenes with backgrounds versus scenes without backgrounds. All but one variable (retfix) showed this effect in all five conditions; background-present scenes had higher values than background-absent scenes in all variables except mfd and rettime, which (necessarily) had lower values with background-present scenes.

In order to determine whether background made the differences between conditions in block 5 greater or less, a separate repeated measures ANOVA was performed on the eye movement variables for block 5 in Experiment 2; we performed this separate analysis since the first analyses using block 5 data also included block 4 data for comparison, which could not also be done for the data in Experiment 1. Table 2.9 provides the ANOVA table and Table 2.6 provides the F values from the contrasts performed between conditions for all variables with significant effects in the ANOVA table (Table 2.9). If one compares these tables to those derived for Experiment 1 (Tables 1.3 and 1.4), it is clear that the differences between conditions are more pronounced with backgrounds present, as is true with the familiarity ratings. Figures 2.16 through 2.29 clearly demonstrate how conditions can be distinguished from each other with background present but not with backgrounds absent.

These differences in eye movement variables between experiments are likely due to the greater informativeness and/or size of the scenes with backgrounds present. Clearly, as the amount of information present in the scene declines, the potential for eye movement differences between observers with and without prior exposure will decline; the more information present, the more potential ways there are to examine the scene.

Discussion

As in Experiment 1, the present experiment provides results that support the argument that object representations are view-dependent. As in Experiment 1, familiarity ratings for scenes changed across exposures in a perceptually specific manner, consistent with the presence of view-dependent scene representations. Specifically, familiarity for scenes shown repeatedly in the same view increases faster than for scenes shown in different views, and most importantly, repeated exposure to scenes in the same view results in lower familiarity when that scene is displayed in a new view (same-different condition) than if the initial exposures had been with different views all along (different-different condition). Since the number of exposures to the scene were equivalent, this suggests that recognition of a scene or object from a new view is dependent on the view “coverage” of the representation built up for that scene. This study showed that this is true whether the stimulus is a complex scene or an object.

This same pattern of results were obtained with the eye movement measures. Almost every eye movement variable showed perceptually specific effects and correlated highly with the direct measure’s results (see Table B.2 in Appendix B). Specifically, perceptually specific effects emerged in the first four blocks where most indirect measures showed a more rapid change in values in the same condition compared to the different condition. In blocks 4 and 5, the primary eye movement variables showed greater perceptual specificity in the same-different condition than in the different-different condition.

The removal of the background in this study had very little effect on the familiarity ratings, but it did have a significant effect on the eye movements. Eye movements with backgrounds present differentiated between conditions better than when backgrounds were not present. This effect has little to do with the issue of scene representation, however, and is most likely driven by the fact that there is less to examine when the backgrounds are removed.

This lack of difference between experiments suggests that outdoor scenes composed of immovable objects are not represented as complex objects but rather as the sum of all (or some) object representations within the scene. Representing the scene as objects with spatial relations between them would seem a more effective strategy than representing the scene as a complex object for such an object would be transparent, a quality very different than normal objects.

In this experiment, we showed that immovable outdoor objects are represented in a view-dependent manner, characteristically similar to the representations of complex outdoor scenes. We also showed that rotated complex scenes are treated the same as rotated objects and suggest that complex scenes are represented as a collection of the representations of their object parts with greater weight placed on the foreground objects over the background. In experiment 3, we test this hypothesis by manipulating the stimuli in such a way that is opposite to the manipulations in experiment 2; we change the scene as a whole but not the primary object within the scene.

Experiment 3: Scene Translation

The experiments described above investigated the nature of the representations constructed over time for naturalistic scenes. In particular, they investigated the separate contributions of changing the scene as a whole or only the primary object within the scene. The following experiment addresses this issue in greater detail by examining the perceptual specificity of scene representations for changes in left-right frame translation. With left-right frame translation, we are changing the scene as a whole but not the primary foreground object. By doing this, we can test the hypothesis that the similar pattern of results in Experiments 1 and 2 was due to the primary object in a scene being given more weight in the scene's representation. If we find that translation has no effect, then we conclude that the primary object is given more weight, but if we find it does have an effect, then we cannot make this claim.

Translation of a scene can also be conceptualized as simply addressing the issue of frame-specific representations: Are the representations specific for the location of objects in the scene relative to the frame bounding it? Research by Intraub and Richardson (1989) and Intraub, et al., (1996) has consistently found that subjects' direct memories for a scene are more "wide-angled" than the actual scene to which they were exposed. When asked to draw scenes that they were previously shown, subjects would consistently draw the scene with the same or wider angle than the actual scene; and when shown other images of the same scene, with the same, wider, or narrower camera angle, subjects would rarely choose the narrower image as the one they saw previously. Intraub suggests that this effect is due to the scene being schematized for its contents, which is to say that the memories are not frame-specific. Intraub, however, never showed images that were translated horizontally from their original, which would entail both the addition and subtraction of scene features, rather than simply all addition, or all subtraction.

In terms of the sensitivity of indirect measures for horizontal translation, Biederman and Cooper (1991) investigated the effects of horizontally translating objects in the visual field and found no effects of horizontal translation on priming. While both the Biederman and Cooper study and the present study shift the position of an object in the visual field, one difference between the studies that makes them difficult to compare is that the present study includes the addition and deletion of peripheral objects while the Biederman and Cooper study simply shifted one object. In addition, the main object in the present study shifts considerably less than it did in the Biederman and Cooper research.

In addition to examining the perceptual specificity of scene representations for translation, this study also examines the effect of amount of stimulus change (2 levels of translation), just as Experiment 1 found that familiarity for new scenes seen previously in only one view decreased in relation to the amount of rotation from the training view. By including two levels of stimulus translation, we can infer the “width” and sensitivity of the scene representation.

We also examine the effect of the subject’s task (scene versus picture recognition) in this study. We do this because the research by Intraub has used a picture recognition task while Experiments 1 and 2 used a scene recognition task; by including both tasks in this study we can make comparisons with both sets of research.

In this study, we examine the perceptual specificity of one direct measure (recognition) and one indirect measure (eye movements) for changes in frame translation of naturalistic scenes. We use recognition as our direct measure in this (and the last) experiment because previous research on perceptual specificity has used this measure and because it is a “cleaner” direct measure of memory than familiarity is from the standpoint of researchers’ interpretation of the data.

By having conditions that either keep the image the same or shift them varying amounts of camera direction (translation), the nature of the scene representations mediating performance on these measures can be inferred. To the extent that the translation condition, which can be conceptualized as the scene changing with the primary object remaining the same, is opposite of the different condition in Experiment 2, where the primary object changed but the “scene” remained the same, we can test our hypothesis that the weight of the representation of complex scenes is on the representation of the primary object in that scene. At another level of conceptualization, we are simply testing our hypothesis that scene representations are frame-specific.

Method

Stimuli. The stimuli for this experiment consist of the same 45 full color scenes used in Experiment 1, plus 3 additional scenes of similar content for a total of 48 target scenes. In addition, there were 48 distractor scenes whose content was equated with those of the target scenes and which appeared in the experiment only once. Of these 48 distractor scenes, 24 appear in block 1 and the other 24 appear in block 2, counterbalanced across subjects. Though the target scenes in this experiment consist of the same scenes in Experiment 1, they differ in the following manner. Each scene has 5 possible images associated with it, the center one being the middle image of the the set of 5 images used in Experiment 1. The other 4 images are not of different views but rather are translated such that the frames of the pictures move while the viewing angle on the primary object remains the same (see Figure 3.1). These images were collected by standing in one location and panning the camera from the left edge of the reference image to its right edge. From these pans, two pairs of images were digitized using the AVID program for Macintosh such that the first images to the left or right of the reference image shared 86% (on average) of the pixels with the reference image while the second set of images to the left or right of the reference image were approximately 59% (on average) similar to the reference image (see Figure 3.2). The similarity of the images was determined by dividing the number of shared horizontal pixels by the total pixel length of the reference image. An effort was made to ensure that the translated image did not differ from the reference image either by deleting a significant portion of the primary object or by adding a new object that seemed more “primary” than the primary object in the reference image.

The design of this study consists of 3 blocks, each with 48 images. The first two blocks are study blocks in which only target reference images and distractor images are shown. The same reference images are shown in both blocks while different distractor images are shown in each block; they are displayed together in a random order. The purpose of showing the same reference images in both blocks is to ensure that the scenes become familiar enough to the subject so that when a change is finally made in block 3, the effect they may have on eye movement variables will be detectable. The ANOVA analyses in Experiment 2 suggest that exposure effects are pronounced enough by block 3.

In block 3, there are 3 conditions: same, translate, and new. In the same condition, subjects see 12 of the same reference pictures seen in the previous blocks. In the translate condition, 12 translated images are used, half of them at 86% similarity (T1) and half of them at 59% similarity (T2), counterbalanced within subjects for right and left of the reference image.

Participants. 32 students and staff (20 females, 12 males) at the University of Illinois were paid 5 dollars for participating in this study. All subjects reported normal or corrected-to-normal vision.

Procedures. Subjects were instructed to examine each scene and make a judgement about how confident they were on a 1-7 scale that they had or had not seen the general scene (scene task, 16 subjects) or exact picture (picture task, 16 subjects) before. A “1” represented “highly confident that I have not seen the scene/picture before”; a “4” represented “uncertain”; a “7” represented “highly confident that I have seen the scene/picture before.” Subjects were informed that scenes may appear translated after the first block of trials. Specifically, subjects in the scene task were told the following:

> During the first set of pictures, the scenes will be unfamiliar to you. In the second and third set of pictures, some scenes will reappear exactly as you saw them before. Others will be completely new. Still others will be of scenes that you have seen

> before but ill be translated (as if the camera shifted a little to the left or right between shots). If you notice that the scene you are seeing is one that you saw before but is translated/shifted from its original version, you should ignore this and make your judgement as if it was not translated. In other words, make your judgement on the basis of the general SCENE and not the specific PICTURE itself. So if you are very certain that you have seen that scene before, even if it is shifted now, you would rate it as a "7."

Subjects in the picture task were told to treat a translated picture as a new picture, to make their judgement on the basis of the specific picture and not the general scene, and were told to rate shifted versions of the same scene with a "1" if they were certain it was shifted. During a practice session of 16 pictures that did not appear in anywhere else in the experiment, the experimenter made sure that subjects understood the instructions. After the practice session, the experiment began. Each trial began with a fixation point in the center of the screen, followed by the stimulus, which remained for 5 seconds, followed by a prompt to make a verbal confidence judgement. In order to eliminate primacy and recency effects, each block contained 2 scene images prior to and after each block of trials. These scenes served only as primacy and recency buffers for all subjects and were not analyzed. Subjects received a short break between blocks.

Hypotheses

The present experiment primarily addresses the issue of the perceptual specificity of scene representations for horizontal translation, as well as the weight of the primary object in a scene's representation. Unlike previous experiments, we expect to not falsify the null hypothesis using either measure and thereby expect to find no significant difference between the same and translated conditions. We expect this lack of an effect because translation-like changes did not affect direct memory measures in Intraub's research, nor did it affect indirect memory measures in research by Biederman and Cooper.   In addition, we do not expect translation to have an effect because we hypothesize that primary objects, which do not change in the translation condition, are weighted more heavily in the complex scene representations, so as long as the primary objects do not change, the "scene" will not change in the subject's mind. If, however, we do falsify the null hypothesis by finding a difference between same and translated (T1 and T2 together), we can test the degree of specificity of these representations by determining whether significant differences emerge between T1 and same and between T1 and T2. If significant differences emerge between these sub-conditions, then we conclude that scene representations are highly perceptually specific for translation.

The second issue addressed in this experiment concerns the effect of task (scene versus picture recognition) on the indirect measure. We want to determine if task has an effect when the scene as a whole changes but no objects within the scene rotate with respect to the viewer. If we falsify the null hypothesis and find a significant main effect of task, or if we find a significant interaction between task and condition, then we conclude that the subject's specific task (scene versus picture recognition) plays an important role in eye movements when a scene must be recognized across different pictures of that scene.

Results

The purpose of this experiment was to investigate the perceptual specificity of scene representations for changes in horizontal translation of the frame of the picture, and to determine whether the subject's task (scene versus picture recognition) has any effect on these results. One direct dependent measure (confidence rating) and three indirect dependent measures (nfix, h2t, and nclust) were recorded.

The analyses in this section include the following: a repeated measures ANOVA using the full design, as well as separate ANOVAs for each of the task groups.

To examine the effect of task (scene versus picture recognition), as well as the perceptual specificity of each of our dependent measures, a repeated measures ANOVA was performed on each of the 4 dependent variables, with condition and block as the 2 within-subject independent variables and task as the between-subjects independent variable (see Table 3.1). These analyses yielded significant effects for all main effects and interaction effects for all 4 dependent variables. Figures 3.3 through 3.6 illustrate that differences between the same and translate conditions arise in the direct measure only in the picture recognition task and in the indirect measures only in the scene recognition task. Such a result demonstrates that the subjects' task has a significant effect on the pattern of results that emerge. Given this, each of the two sets of task data (scene and picture) was analyzed separately below.

<u>Scene recognition task.</u> To determine whether the direct measure showed perceptual specificity for frame translation when the subject was in the scene recognition task condition, a repeated measures ANOVA was performed, with condition and block as the two within-

subject dependent measures. This analysis yielded a main effect of condition, block, and a condition by block interaction. See Table 3.2 for the ANOVA table and the left side of Figure 3.3 for a graphic illustration. Planned comparisons were performed to determine whether same images were recognized any better than translated images and this analysis yielded a nonsignificant difference (p=.59). This demonstrates that same images were recognized no better than translated images, which were both recognized as being different than new images. This lack of a difference is likely due to a ceiling effect and will be discussed later.

To determine whether the indirect measures showed perceptual specificity for frame translation when the subject was in the scene recognition task condition, a repeated measures ANOVA was performed again as above for nfix, h2t, and nclust. For all three variables, significant main effects were found for condition, block, and a condition by task interaction. Table 3.2 provides the ANOVA table and the left side of Figures 3.4, 3.5, and 3.6 illustrate the results for these variables. Planned comparisons were performed between the translation and same data in block 3, as well as translation and new data in block 3 (Table 3.3). Significant effects were found between same and translated data for all three variables. All three variables yielded significant differences between translated images and new images. These analyses indicate that indirect eye movement measures are sensitive to frame translations of a scene.

To further investigate the sensitivity of the indirect measures for frame translations, the translation condition data were split into subgroups T1 (high similarity) and T2 (low similarity). Planned comparisons contrasts were performed between same and T1, same and T2, T1 and T2, and T2 and new. As can be seen in Table 3.4, none of the three variables showed significant effects between same and T1, but same and T2 were significantly different for nfix, h2t, and nclust. T1 and T2 were significantly different for the h2t variable but not for the other variables. Significant differences were found between T2 and new for nfix and h2t but not for nclust. These results demonstrate that the effects of perceptual specificity arise when the scene is translated some amount greater than that between same and T1, and sometimes as small as that between T1 and T2.

<u>Picture recognition task.</u> To determine whether the direct measure showed perceptual specificity for frame translation when the subject was in the picture recognition task group, a repeated measures ANOVA was performed, with condition and block as the two within-subject dependent measures. As in the scene task, this analysis yielded a main effect of condition, block, and a condition by block interaction (see the Table 3.5 and the right side of Figure 3.3). A planned comparison was performed to determine whether same images were recognized any better than translated images and, in contrast to the scene task, this analysis yielded a significant difference [F(1,15)= 272, p<.0001]. This demonstrates that subjects were more certain about having seen same images before compared to translated images for which subjects were completely uncertain (mean=3.9).

To further investigate the perceptual specificity for horizontal frame translation, planned comparisons were performed between same and T1, same and T2, T1 and T2, and T2 and new. Significant differences were found between same and T1 [F(1,15)=15.4, p<.001], T1 and T2 [F(1,15)=96.7, p<.0001], and T2 and new [F(1,15)=18.1, p<.0001]. Such effects are more specific than what was found in this analysis on the indirect measures for the scene recognition group. The similarity of these results is likely due to both measures being driven by the same representations; the differences between the measures is likely a result of the lesser statistical variability in the direct measure.

To determine whether the perceptual specificity for frame translation could be shown with the indirect measures when the subject was in the picture recognition task group, repeated measures ANOVAs were performed again as above for nfix, h2t, and nclust. For all three variables, no main effects of condition or condition by block interaction were found (see Table 3.5 and the right side of Figures 3.4, 3.5, and 3.6). A main effect of block was found for h2t and nclust. These analyses indicate that indirect eye movement measures are only sensitive to frame translations of a scene when the subject's task is to recognize the general scene but not the exact picture. This lack of an effect for the picture task is probably due to the difficulty of the task, forcing subjects to engage in an detailed examination of the picture whether the picture is the same, translated, or new.

Discussion

The purpose of this experiment was to investigate the perceptual specificity of scene representations for horizontal translations of the frame of a scene relative to the scene's primary object, and to test our hypothesis derived from Experiment 2 results that the primary object is weighted more heavily in a scene's representation. In addition, we were also interested in whether the subjects' task would have an effect on these results; so we asked half of the subjects to make recognition judgements on the basis of having seen the same general scene before and the other half of subjects on the basis of having seen the exact picture before.

We found that type of task had a significant influence on both the direct and indirect measures in this experiment. When subjects were asked to make recognition judgements on the basis of the general scene, we found that subjects were no better at recognizing the exact same scene a second time than a translated version of that scene. Thus, we found that the direct measure showed no perceptual specificity for changes in frame translation. This result is different than what we found in Experiments 1 and 2 in which a change in rotation did show a perceptually specific effect. This difference between experiments, however, is likely due to a ceiling effect. Given the perceptually specific effects discussed below in the indirect measures, it is possible that with less exposure a perceptually specific effect would have emerged.

While the direct measure showed no perceptual specificity for translation in the scene recognition task, the indirect measures did show a perceptually specific effect. All indirect variables examined demonstrated this effect. The specificity for frame translation fell somewhere around 67% similarity between study and test image, as significant differences were detected between the same condition and the translated condition, which was composed of images sharing 86% and 59% similarity with the original image.

The perceptual specificity of the indirect measures for frame translation in the scene recognition task was also found for the direct measures in the picture recognition task. When subjects were asked to make recognition judgements on the basis of having seen the exact picture before, subjects were far more certain about pictures that had not translated and were completely uncertain about scenes that had translated.

While the indirect measures in the scene recognition task showed perceptual specificity for translation, the indirect measures in the picture recognition task showed no perceptual specificity for translation. Eye movement measures did not change significantly across conditions or across blocks in the picture recognition group. Such a result creates a pattern of results in which the direct and indirect measures are doubly dissociated across the type of task the subjects were performing. While we found perceptually specific results in the direct measure in the picture task and the indirect measure of the scene task, we found no perceptually specific results in the direct measure of the scene task or the indirect measure of the picture task. Despite this double dissociation pattern, we do not believe that this result reflects the existence of separate representations, one that is perceptually specific and one that is not. Instead, we believe that the results of this experiment are best described as being perceptually specific. In those cases in which we found no perceptual specificity, other factors influenced the results. The perceptual nonspecificity of the indirect variables in the picture condition most likely emerged because the task is simply quite difficult and subjects scan all pictures in a similar manner since they can never be certain that the picture did not shift a small amount. The perceptual nonspecificity of the direct measures in the scene condition most likely emerged due to a ceiling effect of familiarity with the scenes since the mean of both conditions were 6.9 (out of 7.0). Therefore, we conclude that, contrary to our original hypotheses, perceptually specific effects emerged in both the direct measure (in both tasks, though masked in the scene task) and the indirect measures (in the scene task).

In this experiment we found evidence that scene representations are sensitive to the relationship between the frame and the objects in the picture or the relationship between the viewer and the objects in the picture. Our experiment does not discriminate between these two options but the insensitivity for frame position in Intraub's research suggests that subjects are sensitive to the position of the objects within a scene, relative to them.

The perceptually specific effects that emerged in this experiment were contrary to what we had hypothesized given prior research by Intraub and our hypotheses concerning the relatively greater weight allotted to primary objects in a complex scene's representation. In this experiment, where primary objects did not change, we nevertheless found significant effects, suggesting that the primary objects do not enjoy a special status within a complex scene representation. The representation of a complex scene may be such that it is sensitive to change to any quality of the scene. In Experiment 4, we examine this by changing the left-right orientation of the scene, thereby changing the picture without making any changes to the objects themselves or the scene as a whole.

Experiment 4: Scene Reflection

The last issue that we examined is that of orientation-specificity: Are scene representations perceptually specific for left-right orientation (mirror images)? Are representations sensitive to any type of picture change, even when the objects and the scene as a whole does not change? Previous research by Biederman and Cooper (1991) has found that priming tasks do not show perceptual specificity for left-right reflections of objects, but research by scenes Barlett, Gernsbacher, & Till (1987), Bartlett, Till, Gernsbacher, & Norman (1983), and Bartlett, Till, & Levy (1980) has found that direct tests of memory are sensitive to such reflections. As these experiments utilized objects as stimuli, it is uncertain whether such results apply to pictures of complex scenes. Research by Frederickson and Bartlett (1987) that used a direct measure of memory suggests that subject are sensitive to scene reflections but we examine this ourselves in order to allow a better comparison with the previous 3 experiments.

Method

Stimuli. The stimuli in this experiment consist of the same 48 reference images and 48 distractor images used in Experiment 3, except that in the present study, all 96 of these images may appear as target items. Each image also has a corresponding reflected image, which was made simply by flipping the image around the vertical axis in Adobe Photoshop. All other aspects of the design are identical to that of Experiment 3, simply replacing the translation condition with a reflection condition. The assignment of scenes to the same, reflection, and new conditions is completely counterbalanced across subjects such that each scene appears in each condition an equal number of times.

Participants. 24 students and staff (16 females, 8 males) at the University of Illinois were paid 5 dollars for participating in this experiment. All subjects reported normal or corrected-to-normal vision.

Procedures. The procedures for this experiment are identical to that of Experiment 3, except that subjects were informed of and shown examples of scene reflection, and all subjects made “old/old” recognition judgements on the basis of the general scene (not of the exact picture). We use the general scene task because it allows us to compare the results with those acquired in Experiments 1 and 2.

Hypotheses

This experiment only addresses the issue of the perceptual specificity of scene representations for changes in left-right orientation. We expect to falsify the null hypothesis and find a significant difference between the same condition and the reflect condition with both measures.

Results

The purpose of this experiment was to investigate the perceptual specificity of scene representations for changes in the left-right orientation (reflection) of the picture. Unlike Experiment 3, no effect of task was investigated in this experiment. Confidence ratings (direct measure) and a subset of the eye movement measures used in the previous 3 experiments (nfix, h2t, nclust, mfd, leftfix, lefttime, retfix, rettime) were recorded.

To determine whether the direct measure showed perceptual specificity for reflection, a repeated measures ANOVA was performed, with condition and block as the two within-subject dependent measures. This analysis yielded a main effect of condition, block, and a condition by block interaction (see Table 4.1 and Figure 4.1). Planned comparisons were performed to determine whether same images were recognized any better than reflected images and this analysis yielded a nonsignificant difference (p=.59). This demonstrates that same images were recognized no better than reflected images, which were both recognized as being different than new images (Table 4.2).

To determine whether the indirect measures showed perceptual specificity for reflection, repeated measures ANOVAs were performed again as above for nfix, h2t, and nclust (see Table 4.1 and Figures 4.2, 4.3, and 4.4). For nfix, a main effect of condition, block, and a condition by block interaction were found. For h2t, a main effect of condition, block, and a marginal condition by block interaction were found. For nclust, main effects of condition and block were found with no interaction effects. It is worth noting that these effects were not as robust as they were in Experiment 4, despite the higher n.

Planned comparisons were performed between the reflection and same conditions in block 3, as well as reflection and new data in block 3 (see Table 4.2). Significant effects were found between the same and reflected conditions for nfix, h2t (marginally), but not nclust. Significant differences were found between reflected and new conditions for nfix, h2t, and nclust. These analyses indicate that indirect eye movement measures are sensitive to reflections of a scene, though perhaps not as sensitive as they are to frame translations.

To examine whether eye movement measures are more sensitive to the frame translations in Experiment 3 than the reflections in Experiment 4, 2 (Experiment) x 3 (Condition) ANOVAs were conducted on the eye movement measures, the results of which are provided in Table 4.3. All 3 variables showed significant main effects such that the values of all three conditions (same, translate/reflect, new) in Experiment 3 were less than the values of those conditions in Experiment 4. Thus, the values in Experiment 4 were shifted in the more “familiar” direction. Significant interaction effects were also found for all three variables, suggesting greater differences between conditions in Experiment 3 than Experiment 4. T-tests were performed on each of the three conditions between experiments for the nfix and h2t variables, since these are the only two variables that showed significant effects for reflection. For the nfix variable, significant differences were found for the same ($t=6.6$, $p<.01$) and view ($t=4.2$, $p<.05$) conditions, but not the control condition ($t=1.2$, $p=.29$). For the h2t variable, significant differences were found for the same ($t=6.1$, $p<.02$) and view ($t=10.0$, $p<.01$) conditions, but not the control condition ($t=1.5$, $p=.23$). These results suggest that eye movement measures are more sensitive to the frame translations of Experiment 3 than the reflections in Experiment 4, although it is difficult to explain why the same conditions were different between experiments.

Discussion

The purpose of this study was to determine whether scene representations show perceptual specificity for changes in left-right orientation (reflection) of pictures and to determine whether scene representations are sensitive to changes in a picture even when the objects and scene as a whole does not change. The effect of subject task was not investigated in this experiment; all 24 subjects were asked to make an “old/old” recognition judgement about the general scene, regardless of orientation.

The direct memory measure revealed no recognition advantage for seeing the exact same picture than for seeing the reflected version of that picture at test. Therefore, this suggests that the direct memory measure shows no perceptual specificity for left-right orientation just as it showed no specificity for frame translation when the subject’s task was that of general scene recognition. But like Experiment 3, we suggest that this nonspecific effect is due to a ceiling effect since previous research has found perceptually specific effects when the subjects were asked to make recognition judgements on the basis of the specific picture; in addition, subjects in this experiment would often casually comment about a picture being reflected, hence revealing the specificity of their explicit knowledge. If this were not a ceiling effect, however, then this result was probably mediated by orientation-invariant representations, since previous research has found evidence for both orientation-dependent and orientation-invariant representations.

The eye movement measures did show significant differences between pictures seen in the same orientation and those seen in the reverse orientation. Thus, we find that while the direct measures show no perceptual specificity, the indirect measures nevertheless do reveal the perceptual specificity of the underlying scene representations. Since perceptually specific effects emerged even though neither the individual object nor the scene as a whole changed, I suggest that complex scene representations are sensitive to a variety of changes to a scene, including viewer-object spatial relations.

General Discussion

The primary contribution of this thesis is that it provides one of the first glimpses into the nature of complex scene representations. Until very recently, little research had been conducted on this topic. Instead, most research has focused on how individual objects are represented. This thesis extends the research on object representation into the arena of complex scenes and examines the nature of scene representations by determining the sensitivity of these representations for various types of changes to the scene. In Experiment 1, we examined the sensitivity of complex scene representations to rotations in depth and through a variety of manipulations attempted to determine the angular breadth of these representations. From our results, we make the following two claims. First, we claim that scene representations, like object representations, have view dependent components. Second, we claim that these representations cover a breadth of at least 45 degrees and have a specificity of 15 degrees or less.

We suggest that scene representations have view dependent components for the following two reasons. First, scenes displayed repeatedly in the same view increase in familiarity at a faster rate and elicit a different set of eye movements than scenes displayed repeatedly in a different view. Secondly, this sensitivity to view also emerged in the last block of Experiment 1 where a new view was displayed of a scene that was shown repeatedly in the same view. Both familiarity ratings and eye movements were sensitive to this change in the last block and a closer analysis of the familiarity ratings in that condition demonstrated that familiarity decreased as a linear function of the degree rotation from the learned view.

This pattern of results is most often interpreted as evidence for a viewer-centered representation (Hayward and Tarr, 1997).

Results like these were also obtained by Diwadkar and McNamara (1997) although their stimuli consisted of movable indoor objects and each scene differed from other scenes only by the spatial arrangement of the same 6 objects. View dependent effects have also been obtained in electrophysiological studies (Logothetis & Pauls, 1995; Logothetis, et al., 1994, 1995). In these studies, cells were found that had firing rates that were preferential to particular views of objects and depolarized as a linear function of the degree of angular rotation from the nearest learned view. In addition, they found that if sufficient “overlap” of representations occurred (less than 40 degrees of rotation), view-invariant firing rates would occur.

Contrary to the viewer-centered theory of representation, Biederman and Gerhardstein (1993) claim that, for objects, rotation from the learned view should not affect performance unless critical parts of the object become occluded or emerge from occlusion. If we apply this reasoning to complex scenes, we must conclude that either we have evidence for view dependent representations (or representations with view dependent components), or critical parts of the scene changed across rotations. We concede, for two reasons, that we cannot claim with absolute certainty that critical parts of the scene did not become occluded or emerge from occlusion. First, we cannot be certain because it is unknown exactly what the critical parts of a scene are, and we cannot control for that which we are not knowledgeable. Second, the amount of control applied during the stimulus construction process to ensure that critical parts remained visible across rotation was not as strict as it is (and is required) in experiments that have attempted to refute Biederman and Gerhardstein's (1993) hypothesis. Thus we cannot entirely rule out an abstract representation hypothesis. Nevertheless, a great deal of care was taken to ensure that across rotation, the picture maintained the subjective quality of representing the same scene; no foreground objects ever occluded the scene and the critical parts of the primary object were most often visible from all viewpoints. Furthermore, since differences in familiarity ratings between the same condition and the different condition were maintained up until block 5, where the maximum angular rotation from a learned view in the different condition could be only 15 degrees, Biederman's hypothesis would only apply if critical parts of scenes were frequently occluded by rotations of only 15 degrees. This is very unlikely. For these reasons, and because the pattern of results so closely resembles previous research results in support of view dependent representations of objects, we conclude that scene representations, like object representations, have view dependent components.

We claim that these (view dependent) representations cover an angular breadth of at least 45 degrees and are sensitive to rotations of as little as 15 degrees. We claim this because subjects, while showing decreased familiarity with 45-degree rotated scenes in the same-different condition in block 5, were nonetheless able to recognize 45-degree rotated scenes, rating them significantly higher in familiarity than control scenes. Extrapolation of the familiarity ratings for the 3 rotation levels of this condition suggests that the breadth of the representations is much wider. This would be inconsistent with object recognition research on rhesus monkeys by Logothetis and Pauls (1995) which finds that breadth of representation is no more than 40 degrees, but consistent with the object recognition research on humans by Srinivas (1995), which has found that breadth of representation can be up to 67 degrees. The analysis of these levels of rotation in the same-different condition showed significant effects for rotations of 30 degrees or higher and marginally significant effects for rotations of 15 degrees. It is noteworthy that while the 15 degree rotation level of the same-different condition was not significantly different from the different-different condition (which is, in effect, a 15 degree rotation condition also), the different-different condition, but not the 15 degree level of the same-different condition, was signficantly different from the 30 degree rotation level of the same-different condition. Presumably this is due to the view in the different-different condition being most often “surrounded” by broadly-tuned, overlapping representations while the view in the same-different condition having only one broadly-tuned, overlapping representation nearby; this difference is what Bulthoff, Edelman and Tarr (1995) refer to as interpolation (different-different condition) versus extrapolation (for the same-different condition). Further evidence for 15 degree sensitivity came from the experiment reported in Appendix D, where only rotations of 15 and 30 degrees occurred and subjects were given less exposure to the scenes. In this experiment we found that 15 degree rotations were significantly different from the nonrotated scenes, in our direct measure. So with less exposure the maximum sensitivity of these representations for rotations in depth can be better ascertained. Future research might focus on how sensitivity and breadth of scene representations varies with the amount of exposure.

Given the similarity of our results in Experiment 1 to the object representation literature, we conducted Experiment 2 in order to determine the extent to which our Experiment 1 effects were being primarily mediated by the representation of the central/primary object in the scene. This experiment was identical to Experiment 1 except that we removed the backgrounds from the Experiment 1 stimuli such that only the primary object remained. This experiment yielded two important results.

First, we found that removing the background significantly reduced the sensitivity of the eye movement measures in block 5 to different exposure conditions for while discriminations between conditions were possible in Experiment 1, they were not possible in Experiment 2. This effect, we suggest, has little to do with the issue of object and scene representations and more to do with simply having less to look at with backgrounds absent. With less to look at, and hence less "gist" to be derived, the eye movement variables have fewer degrees of freedom, must shift toward the "familiar" direction, and become less sensitive.

Second, we found that removing the background had virtually no effect on the overall pattern of results in the direct measure. Since the absence of other objects in the scene did not affect the pattern of results, it seems that either the primary object in these scenes carries most of the weight in the scene representation or the characteristics of scene representations are identical to the characteristics of the representations of the objects that comprise these scenes. Such an interpretation is consistent with previous research that has concluded that scenes are organized in a hierarchical manner, with features that are functionally more significant given greater priority (McNamara, 1986; Taylor & Tversky, 1992).

To test this hypothesis, as well as the basis for Intraub's boundary-expansion effect, we conducted Experiment 3, which determined the sensitivity of scene representations to horizontal translation of the frame of the scene relative to the objects in it. By performing this manipulation, we were able to make changes to the scene without making any changes to the primary object. If the primary object truly carries the weight of the scene representation, then this manipulation should have no effect on our measures. If this manipulation has an effect, then we would conclude that the primary object does not carry the weight in the scene representation (i.e. the primary object is not the sole determinant of the results of the direct and indirect measures).

We found in Experiment 3 that subjects are sensitive to translation. Recognition performance showed this effect when subjects were asked if they had seen the exact picture before, and eye movement performance showed this effect when subjects were asked if they had seen the general scene before. This result has two important implications. First, the significant effects suggest that the primary object of the scene does not carry all of the weight of the scene's representation, as the results of Experiments 1 and 2 might have led one to think. This is so because by translating the scene, we changed the scene without changing the primary object, which is the opposite of what we did in Experiment 2. Since manipulating the picture in both experiments yielded significant effects in both measures, we cannot conclude that the scene is represented simply by representing its primary object. This interpretation is consistent with research by Missal, Vogels, and Orban (1997) in which the responses of neurons in the inferior temporal lobe of macaque monkeys were found to respond not only to foreground shapes but to background shapes as well. Nevertheless, the fact that rotation of the scene or object in the 5th block of Experiments 1 and 2 resulted in significant differences between rotated and same conditions while translation showed no such effects in the 3rd block of Experiment 3 suggests that while the primary object may not be the only relevant aspect of the scene, it is given more weight in the scene's representation. Whether this difference is statistically significant cannot be determined, however, because it would require comparison of results obtained at different levels of exposure and because the familarity ratings cannot be compared to the confidence ratings.

The second implication to be drawn from Experiment 3 is that it is not the case that Intraub's boundary-expansion effects were due to scene representations simply being frame invariant. Instead, these frame-specific results are consistent with Intraub's boundary-expansion explanation. The results from Experiment 4 suggest that the sensitivity for frame-translation is due to primarily to the deletion of a portion of the scene but also to the change in the spatial relationship between the primary object and the frame (or the viewer), as will be explained below.

Finally, we examined the level of specificity of the representations for frame translation by using pictures that were shifted two different distances from the reference image. As both of our measures clearly distinguished between the same condition and the translate condition (an average of 67.5% pixel overlap), we examined whether our measures could distinguish between the reference and the high similarity image (86% shift) or the high similarity image and the low similarity image (59% shift) to determine the specificity of the measures. In addition, we also determined which levels were distinguishable from the control condition in order to determine the width of the representation. We found that the representations were relatively specific since significant differences existed between each of these levels with the direct measure and a subset of the eye movement measures, and we found them to be broad as well, since significant differences were found between the new items and other conditions in the direct measure and some eye movement measures.

In Experiment 4, we examined perceptual specificity for scene orientation in order to attempt to clarify the issue of how much weight the primary object holds in the scene representation and to determine to what aspects of scene translation the representation is sensitive. While previous research has already investigated this topic of orientation specificity, we examined the effect of this manipulation using our stimuli and our dependent measures so that we could more effectively compare the effects of reflection to the effects of our other manipulations in previous experiments. By changing the left-right orientation of a scene, we change the picture without adding or deleting anything to the scene or the primary object in the scene. The only change that occurs is that the relative position of the parts (objects) of the scene change their position relative to the frame (or the viewer). If we find that this manipulation has no effect, this would be consistent with the hypothesis the primary object carries the weight of the scene representation and that the effects in Experiment 3 were due to the deleting of scene parts, not the movement of the primary object relative to the viewer (or frame). But if we find that orientation changes do have an effect wherein recognition performance and eye movements are different for reflected images than nonreflected images, then this suggests that the effects in Experiment 3 may have been due to the deleting of scene parts and/or the movement of the primary object relative to the viewer (or frame).

Unfortunately, the results we obtained in Experiment 4 were somewhat equivocal with respect to whether scene representations are sensitive to changes in orientation.

Orientation did not have an effect on the confidence ratings or the eye movement measures (except nfix). Nevertheless, subjects would often mention during the experiment that they were aware of an orientation change and prior research has shown that direct measures are usually sensitive to changes in orientation (Bartlett, Gernsbacher, & Till, 1987; Bartlett, Till, Gernsbacher, & Norman, 1983; Bartlett, Till, & Levy, 1980). Therefore, I suggest that given the consistent robustness of the nfix variable in the prior 3 experiments, it is reasonable to emphasize the results of this variable over the others and conclude that scene representations are sensitive to orientation changes. If this interpretation is correct, then this suggests that the primary object does not carry the weight of the scene representation since we found an effect without any changes occuring to the object. In other words, parts of the scene other than the primary object are an essential, nondisposable component in the scene's representation. This interpretation is the same one we gave initially above for Experiment 3. In addition, orientation sensitivity suggests that the effects in Experiment 3 may have been a result of deleting scene parts and/or the movement of the primary object relative to the frame (or viewer).

Taken as a whole, these experiments demonstrate that the representations of complex scenes are sensitive to a variety of changes to the scene, similar to, or perhaps even more sensitive than, object representations. Our primary focus was to extend the object representation literature into the realm of scene representation and this we did by showing that scene representations show a similar sensitivity for viewpoint to the sensitivity of object representations. We had also hoped to gain some insight into how multiple objects (foreground and background) and their spatial relations (with each other and with the observer) are dealt with in a scene representation but conflicting results in Experiments 2 and 3 emerged, so future studies suggested in the next section will hopefully elucidate this issue. The biggest issue, which is whether scene and object representations are stuctured in a similar manner, remains unaddressed and is discussed further in the Future Directions section.

Future Directions

The four experiments in this thesis attempted to distinguish between the various components of a scene in an attempt to determine the critical attributes of scene representations. In Experiment 1 we changed the scene, the objects, and the picture by rotating the scene in the depth plane. In Experiment 2, we changed the object and the picture, but not the scene, by rotating a backgroundless object in the depth plane. In Experiment 3 we changed the picture and the scene, but not the primary object, by translating the picture. And in Experiment 4, we changed the picture, but not the objects or the scene, by reflecting the picture. Admittedly, one could argue that the rotating a single object with no background in Experiment 2 is not the best example of changing the object and the picture, but not the scene; furthermore, translating the image in Experiment 3 may not be the best example of changing the picture and the scene, but not the objects. For these two reasons, the following experiments are proposed as means of creating these conditions to better investigate scene representations.

In Experiment 1, the scene and the objects rotated together. In Experiments 2 and 3 we attempted to examine the separate contributions of the scene as a whole and the objects themselves. In order to truly examine the separate contributions of the scene as a whole and the objects themselves, however, we need to design experiments in which we are able to manipulate the objects independently from the scene. In particular we need to design the following two experiments. In the first experiment, the scene would not change with respect to the viewer, but the individual objects would change with respect to the viewer. To do this, we would rotate only the objects while the background stays constant. In a second experiment, we would do the opposite such that the scene would rotate with respect to the viewer but the objects would not rotate with respect to the viewer. In this way, the spatial relationship between the objects and the viewer changes but the surface characteristics of the object remain the same.

Three possible options exist for the type of stimuli that could be used in such a study. Outdoor images such as the ones used in this thesis would not be used since it would not be possible to rotate large objects such as houses. Therefore, if photographs are to be used, they must be of scenes that include movable objects. The second option is to generate these scenes using 3-D computer animation. This is a better alternative since it allows greater control over the stimuli. Objects can be rotated, reflected, and shifted with respect to the viewer while the scene can be independently rotated, reflected, and shifted with the respect to the viewer as well. The third and best option is to generate these images in a virtual reality environment. This option gives us all of the flexibility over the scene's appearance that the computer animation option provides, and it has the additional benefit of being three-dimensional, eliminating the perceptual leap between pictures and real scenes. Research by Sholl and Nolin (1997) on place representation using real scenes has shown that this perceptual leap is significant for they find view dependent effects emerge unless a horizontal viewing angle is used during encoding, a room-sized test space is used, and "on-path" testing is conducted. These criteria can only be met in a real-world or virtual reality environment, but not with a computer monitor.

The design of such studies would be similar to those of Experiments 3 and 4. Subjects would study a set of scenes and then be tested for their recognition of the scenes. Half of the scenes at test would be completely new scenes while a fourth would be the same picture of a scene displayed before and the remaining scenes would be of a changed picture of a scene displayed before. Depending on the experiment, the change would be of either the objects rotating or the scene background rotating relative to the observer. Rotation conditions would have two levels in order to test the sensitivity of the representations.

The dependent measures in these studies would include a verbal report of recognition. If a virtual reality environment is not used, eye movements could be collected again and the verbal report would again be that of confidence ratings. If a virtual reality environment is not used and eye movements are not recorded, a “yes/no” recognition paradigm would be used and reaction times would be recorded, as measures of accuracy and latency are commonly used in the object representation literature. If a virtual reality environment is used, only a verbal report would be taken and reaction times would be recorded if this is possible in that setting. While it would be informative to determine if eyes move towards the parts of the scene that changes relative to the observer or to the objects that move relative to the scene, examining these questions in a three-dimensional space without eye movements would be more informative and hence preferred.

An additional benefit of using a virtual reality environment is that we can explore these issues in real time, rather than with time-suspended pictures. This allows us to expand into two additional paradigms that expand upon the issues addressed here. In the first additional paradigm, virtual rooms can be created in which subjects would have to find an object that they saw in that room before. By performing the same set of stimulus manipulations on the room that were suggested above, we can examine the effect of such manipulations on the time it takes for the subject to perform the task. In this way, such studies are akin to virtual radial arm mazes such as those performed using rats.

Whether these experiments are conducted using computer animation or virtual reality technology, the manipulations discussed above should be performed on two qualitatively different sets of scenes. In one set we would continue with the use of immovable outdoor objects as stimuli while in another set we would use movable indoor objects as stimuli. These stimuli could be used in the same or separate experiments but it is important that they are both used since it is likely that we represent both types differently given the unique ways in which we interact with each of them. Conducting indoor objects versions of the experiments conducted for this thesis would not be necessary as no unusual manipulations were performed. But the proposed studies are likely to show different patterns of results with the different types of scenes given that rotation of objects relative to one another is common for small indoor objects but not large outdoor ones.

To directly address the issue of whether complex scenes are represented as superordinate objects or rather as a collection of their subordinate parts (objects), electrophysiological studies such as those by Puce, et al., (1995) ought to be conducted. Using a similar methodology, the activity of inferior temporal neurons could be recorded in monkeys while they learned to recognize complex scenes. Once neurons were found that responded preferentially to particular scenes, individual components of the scene could be displayed in a rearranged manner. If one found that all of the neurons that responded to a scene responded equally well to the rearranged display, then one would conclude that scenes are not represented as unique superordinate objects. If one found that all cells that responded preferentially to a complex scene showed a significant decrease in activity when the rearranged scene was displayed, then one might conclude that scenes are uniquely represented.

Another neuroscientific approach to this issue would be to examine with functional magnetic resonance imaging (fMRI) whether scene recognition across rotations in depth utilize the same brain regions that object recognition across rotations in depth would utilize. If the between-object spatial relationships of a scene are analogous to the between-feature spatial relationships of an object, then we would expect similar patterns of activity during a task that requires recognition across viewpoints. But to the extent that these two types of spatial relationships are not analogous, then we would expect a different pattern of results. In particular, if they are different one would expect more activity in right parietal regions during the scene recognition task than the object recognition task, which normally utilzes more left temporal lobe processing (Sergent, Ohta, & MacDonald, 1992; Kosslyn, et al., 1994).

The difficulty in performing such a study would be the confound of complexity wherein differences in activity could be attributed to a scene simply being more complex rather than qualitatively different. For this reason, at least two experiments would need to be conducted. In one, we would use scenes and objects such as the ones used in Experiments 1 and 2, respectively; these stimuli do not correct for the confound of complexity but allow comparisons between the types of objects used in previous research and “real world” scenes. In another, we would use “scenes” composed of simple geometric objects separated by a black background and “objects,” which would be similar to the complex spheriod objects used by Logothetis, Pauls, and Poggio (1995); these stimuli may not inform us about “real world” scenes but they will allow us to look at the effect of multiple objects. The results from both studies together will allow us to draw more accurate conclusions than either study would if conducted alone. In both studies, scenes and objects would be displayed one at a time within separate blocks, and subjects would be required to learn the scenes and objects and then recognize them from different views. The design might involve something such as: study object, test object, study scene, test scene; this would be repeated numerous times with an appropriate control condition in between each set. Differences in activity between the scene and object conditions, particularly differences in parietal regions, would suggest that scenes are not represented as superordinate objects but rather as the sum of individual object representations.

Finally, I would also like to conduct another line of experiments that investigates the related topic of perceptual specificity of representations for object motion. Given that the most meaningful object motion we experience is that of animate motion, this study would require the use of stimuli in which humans are observed performing simple motions. A set of movements would be used and a similar number of people would be filmed performing each of these movements. Movements would be performed by members of a dance department since they would be most skilled at making the “same” movement across all individuals. Each dancer would also be filmed making a movement which is similar but slightly different. Subjects would view videotapes of the dancers performing the movements such that each dancer is seen only once and each movement is seen only once. In one experiment we investigate the effects of dancer (old, different, new) and movement (old, new) to determine the extent to which motion recognition is affected by prior exposure to the object performing the motion. In other words, are movements represented separately from the representation of the object moving or are the movement and object representations linked? In this study, subjects would see at test either an “old” dancer performing an “old” movement, an “old” dancer performing a “new” movement, a “old” dancer performing an “old” movement performed at study by a different dancer, a “new” dancer performing an “old” movement, or a “new” dancer performing a “new” movement. Subjects would be asked to state whether they have seen that exact movement before using confidence ratings. In a second experiment, we examine the effects of dancer (old, different) and movement (old, altered, new) to determine the sensitivity of the representations for object motion. During test, subjects would see either an “old” dancer performing an “old” movement, an “old” dancer performing an “altered” movement, an “old” dancer performing a “new” movement, a “different” dancer performing an “old” movement, a “different” dancer performing an “altered” movement, and a “different” dancer performing a “new” movement. As in the first experiment, subjects would make judgements about whether the movement is the same using confidence ratings.

The research discussed in this thesis extends the object representation literature into the arena of scene representation. The virtual reality studies proposed above extend this work even further by using stimuli that have three-dimensional characteristics. The movement studies proposed above extend the object representation literature to encompass the representation of object movement. The combination of the results from these proposed paradigms may provide us with the big picture of how we represent our complex, multiobject world.

Conclusion

In this thesis, we expanded the research on object representation into the arena of complex scene representation. We found in Experiments 1 and 2 that the representations of complex scenes, like the representations of objects, are composed of view-dependent components. In Experiments 3 and 4 we examined the sensitivity of scene representations for translation and reflection in order to determine what components of a scene are critical in its representation. Though we found sensitivity for translation and reflection, the results still suggest that the primary object in the scene is weighted more heavily in a scene's representation. Future research is suggested that can more effectively discern the nature of scene representations.

Table 1.1. Analysis of Variance of Familiarity Ratings

Source	Blocks 1 thru 4		Blocks 4 and 5	
	df	F(MS)	df	F(MS)
Block (B)	3	228 (282)**	1	20.9 (5.6)**
Condition (C)	2	340 (522)**	4	352 (353)**
B x C	6	177 (60)**	4	14.0 (4.4)**
Subject MS	44	(5.1)	44	(6.1)

*p<.05. **p<.01

Table 1.2. Effect of Changing View in the Same-different Condition

Source	Degree Rotation	F
Different-different v. 15 deg rotation	0	0.4
Same-same v. Different-different	15	1.5
Same-same v. 15 deg rotation	15	3.4+
15 deg v. 30 deg rotation	15	2.2
30 deg v. 45 deg rotation	15	1.2
Different-different v. 30 deg rotation	15	4.4*
Same-same v. 30 deg rotation	30	11.1**
15 deg v. 45 deg rotation	30	6.6**

Note: 5 trials per subject in Same-same and Different-different conditions. 2 trials per subject in 15 and 45 degree rotation conditions. 1 trial per subject in 30 degree rotation conditions. +$p<.07$. *$p<.05$. **$p<.01$.

Table 1.3. F Values (MS) for All Variables in Block 5

Source	df	F										
		rating	h1	h1t	h2	h2t	mfd	nclust	nfix	leftfix	retfix	rettime
Condition	4	257**	10.9**	13.3**	13.0**	15.4**	5.4**	16.0**	17.5**	3.7**	0.8	4.5**
		(186)	(1.8)	(2.1)	(1.9)	(2.4)	(15440)	(30.5)	(27.7)	(0.0)	(1.5)	1018840
Subject MS	44	(3.3)	(1.1)	(1.1)	(1.0)	(1.1)	(25820)	(16.4)	(20.0)	(0.1)	(5.3)	(1614820)

*p<.05. **p<.01.

Table 1.4. F Values (MS) from Contrasts on Eye Movement Variables between Most Conditions in Block 5 (n=45)

Conditions	F												
	rating	h1	h1t	h2	h2t	mfd	nclust	nfix	leftfix	rettime	s1	s1t	s2	
C*SS	700**	31.2**	38.7**	43.9**	49.7**	15.9**	44.8**	52.2**	10.6**	13.1**	14.3**	14.2**	12.9**	1
	(508)	(5.0)	(6.1)	(6.5)	(7.7)	(45830)	(85.6)	(82.3)	(0.1)	(2972070)	(0.1)	(0.1)	(0.1)	(
C*SD	472**	8.1**	9.8**	6.2*	6.6*	6.5*	9.5**	12.1**			7.8**	7.1**		
	(343)	(1.3)	(1.5)	(0.9)	(1.0)	(18700)	(18.1)	(19.0)			(0.1)	(0.1)		
C*DD	636**	11.5**	14.4**	15.2**	17.5**	15.4**	18.8**	28.0**		5.3*	7.8**	7.3**	6.0*	4
	(461)	(11.9)	(2.3)	(2.3)	(2.7)	(44290)	(36.0)	(44.1)		(1193130)	(0.1)	(0.1)	(0.0)	(
SS*SD	22.3**	7.5**	9.6**	17.2**	20.0**		13.0**	14.0**	5.3*	11.4**				
	(16.2)	(1.2)	(1.5)	(2.5)	(3.1)		(24.9)	(22.2)	(0.0)	(2569200)				
SS*DD		4.8*	5.9*	7.4**	8.3**		5.5*							
		(0.8)	(0.9)	(1.1)	(1.3)		(10.6)							
SD*DD	12.2**									4.2*				
	(8.8)									(943250)				

Note: Only values reaching significance ($p<.05$) listed.

S=Same, D=Different, C=Control

*$p<.05$, **$p<.01$

Table 2.1. ANOVA Table for Selected Variables in Experiment 2 (n=44)

Source	df	F			
		rating(+)	nfix	h2t	nclust
Blocks 1 through 4					
Condition (C)	2	948 (636)**	26.9 (47.3)**	14.4 (3.4)**	9.0 (18.5)**
Block (B)	3	772 (360)**	7.4 (24.1)**	11.4 (1.8)**	6.4 (14.2)**
C x B	6	315 (78)**	6.5 (5.6)**	5.7 (0.5)**	3.7 (3.6)**
Subject MS	43	(1.4)	(17.7)	(0.6)	(9.6)
Blocks 4 and 5					
Condition (C)	4	1200 (437)**	20.34 (1.0)**	10.4 (0.2)**	6.3 (3.7)**
Block (B)	1	1.6 (0.3)	0.6 (41.1)	1.4 (2.9)	2.5 (16.1)
C x B	4	11.7 (2.0)**	8.7 (8.5)**	5.7 (0.6)**	1.5 (1.5)
Subject MS	43	(1.2)	(23.0)	(1.0)	(11.5)
Block 5 (+)					
Condition	4	–	19.1 (27.4)**	9.1 (1.8)**	6.1 (11.1)**
Subject MS	44	–	(11.7)	(0.5)	(5.6)

(+) n=45, df=44 for rating

*p<.05. **p<.01

Table 2.2. ANOVA Table for All Variables in First Four Blocks of Experiment 2

Source	df	F (+)												
		rating	h1	h1t	h2	h2t	mfd	nclust	nfix	retfix	rettime	leftfix	s1	s1t	
Condition (C)	2	948**	3.2*	5.3**	10.1**	14.4**	3.7*	9.0**	26.9**	1.2	10.9**	0.1	0.4	0.4	1.
		(636)	(0.5)	(0.9)	(2.4)	(3.4)	(26620)	(18.5)	(47.3)	(1.2)	(2295600)	(0.0)	(0.0)	(0.0)	(0
Block (B)	3	772**	3.1*	4.5**	8.4**	11.4**	1.9	6.4**	7.4**	1.0	7.1**	1.3	0.9	0.9	2.
		(360)	(0.6)	(0.8)	(1.3)	(1.8)	(17260)	(14.2)	(24.1)	(1.7)	(2409100)	(0.0)	(0.0)	(0.0)	(0
C x B	6	315**	2.9**	3.7**	4.1**	5.7**	1.1	3.7**	6.5**	2.9**	5.8**	0.8	1.8	1.7	1.
		(78.0)	(0.3)	(0.3)	(0.4)	(0.5)	(4520)	(3.6)	(5.6)	(2.7)	(888800)	(0.0)	(0.0)	(0.0)	(0
Subject MS	43	(1.4)	(0.8)	(0.8)	(0.6)	(0.6)	(33600)	(9.6)	(17.7)	(5.5)	(1050400)	(0.1)	(0.0)	(0.0)	(0

(+) n=45, df=44 for Familiarity; n=41, df=40 for Path, Fixation, and Return

*p<.05. **p<.01.

Table 2.3. F Values (MS) from Contrasts on All Eye Movement Variables between Conditions in Blocks 2, 3, and 4

Block	Conditions	F						
		h1	h1t	h2	h2t	mfd (+)	nclust	nfi
2	S * C	8.3 (0.7)**	11.3 (1.0)**	9.8 (0.9)**	12.2 (1.2)**		13.5 (13.0)**	11.4 (9.
	D * C							
	S * D				4.8 (0.5)*		4.0 (3.9)*	4.1 (3.6)
3	S * C	8.2 (0.7)**	11.8 (1.0)**	30.4 (2.9)**	42.0 (4.0)**	10.32 (41520)**	22.8 (22.0)**	65.5 (56
	D * C	10.6 (0.9)**	12.6 (1.1)**	9.0 (0.9)**	12.0 (1.1)**		15.3 (14.8)**	13.4 (11
	S * D			6.3 (0.6)*	9.1 (0.9)**	6.12 (24620)*		19.7 (17
4	S * C	4.4 (0.4)*	10.9 (0.9)**	34.2 (3.2)**	48.5 (4.6)**	4.00 (15890)*	18.7 (18.1)**	67.1 (58
	D * C		4.7 (0.4)*	4.9 (0.5)*	7.5 (0.7)**		4.6 (4.4)*	7.1 (6.1)
	S * D			13.3 (1.3)**	17.8 (1.7)**	6.1 (24460)**	4.8 (4.6)*	30.6 (26

Note: Only values reaching significance listed. No significant differences were found between conditions in block 1.

(+) denotes no interaction effects in all five blocks

S=Same, D=Different, C=Control

*p<.05, **p<.01

Table 2.4. ANOVA Table for All Variables in Blocks 4 and 5

Source	df	F												
		rating(+)	h1	h1t	h2	h2t	mfd	nclust	nfix	retfix	rettime	leftfix	s1	s1
Condition (C)	4	1200**	2.5*	4.6**	7.0**	10.4**	2.4	6.3**	20.4**	2.5*	9.6**	1.0	0.4	0.5
		(437)	(0.2)	(0.1)	(0.1)	(0.2)	(4.1)	(3.7)	(1.0)	(0.2)	(58950)	(0.0)	(0.0)	(0.
Block (B)	1	1.6	1.2	0.9	1.0	1.4	0.0	2.5	0.6	0.2	0.2	0.6	0.5	0.1
		(0.3)	(0.6)	(1.0)	(2.0)	(2.9)	(29920)	(16.1)	(41.1)	(3.7)	(2985670)	(0.0)	(0.0)	(0.
C x B	4	11.7**	0.9	0.9	4.0**	5.7**	1.5	1.6	8.7**	4.5**	7.4**	1.3	1.5	1.2
		(2.0)	(0.1)	(0.1)	(0.4)	(0.6)	(20570)	(1.5)	(8.5)	(5.1)	(1749620)	(0.0)	(0.0)	(0.
Subject MS	43	(1.2)	(0.9)	(1.0)	(0.8)	(1.0)	(63080)	(11.5)	(23.0)	(6.6)	(1380580)	(0.2)	(0.1)	(0.

(+) n=45, df=44

*p<.05. **p<.01.

Table 2.5. F Values (MS) from Contrasts on Selected Eye Movement Variables between Conditions within Block 5 (n=45)

Conditions	F(+)		
	nfix	h2t	nclust
C*SS	97.4 (94.5)**	48.2 (5.0)**	34.6 (34.4)**
C*SD		26.6 (2.8)**	25.7 (25.5)**
C*DD	31.7 (30.7)**	25.7 (2.7)**	18.5 (18.4)**
C*DS	58.4 (56.6)**	51.5 (5.4)**	29.5 (29.3)**
SS*SD	22.1 (21.4)**		
SS*DD	18.0 (17.5)**		
SS*DS	5.0 (4.8)*		
SD*DD			
SD*DS	26.8 (26.0)**		
DD*DS	4.1 (3.9)*	4.4 (0.5)*	

Note: Only values reaching significance listed.

(+) Analysis performed using data from blocks 4 and 5.

S=Same, D=Different, C=Control

*p<.05, **p<.01

Table 2.6. F Values from Contrasts on Selected Eye Movement Variables between Conditions within Block 5 (n=45)

Conditions	F(+)							
	h1	h1t	h2	h2t	nclust	nfix	retfix	r
C*SS	6.2 (1.1)*	10.1 (1.7)**	16.7 (3.3)**	26.3 (5.1)**	19.1 (34.8)**	69.3 (99.4)**		22.5 (6
C*SD	8.8 (1.5)**	12.0 (2.0)**	11.2 (2.2)**	14.5 (2.8)**	13.2 (24.1)**	19.1 (27.3)**		
C*DD	7.0 (1.2)**	8.3 (1.4)**	12.2 (2.4)**	15.0 (2.9)**	11.0 (20.2)**	24.3 (34.8)**		
C*DS	6.2 (1.1)*	11.1 (1.9)**	19.8 (4.0)**	28.1 (5.5)**	15.1 (27.6)**	40.8 (58.6)**	5.0 (6.8)*	12.2 (3
SS*SD						15.7 (22.5)**		8.4 (22
SS*DD						11.5 (16.6)**		13.1 (3
SS*DS								
SD*DD							4.8 (6.5)*	
SD*DS						4.1 (5.9)*		
DD*DS							7.1 (9.6)**	5.6 (15

Note: Only values reaching significance (p<.05) listed.

(+) Analysis performed using only block 5 data

S=Same, D=Different, C=Control

*p<.05, **p<.01

Table 2.7. ANOVA Table Comparing Direct Measure in First Four Blocks of Experiments 1 and 2

Source	df	F (MS)
Experiment (E)	1	2.0 (6.4)
Subject MS	88	(3.3)
Condition (C)	2	1047 (1155)**
Block (B)	3	751 (639)**
C x B	6	467 (137)**
C x E	2	4.0 (4.3)*
B x E	3	2.8 (2.4)*
C x B x E	6	2.4 (0.7)*

*$p<.05$. **$p<.01$.

Table 2.8. F Values (MS) for All Variables between Block 5 of Experiments 1 and 2 (n=90)

Source	df	F										
		rating	h1	h1t	h2	h2t	mfd	nclust	nfix	leftfix	retfix	rettime
Background (B)	1	6.3**	18.1**	18.5**	21.6**	20.3**	9.4**	25.3**	12.5**	3.4	5.9*	3.7
		(12.3)	(14.4)	(14.8)	(15.3)	(16.0)	(306950)	(278)	(197)	(0.3)	(27.7)	(4704850)
Condition (C)	4	836**	10.9**	14.8**	17.1**	22.9**	4.9**	19.9**	36.0**	3.9**	1.1	10.3**
		(405)	(1.8)	(2.4)	(3.0)	(4.0)	(45820)	(37.3)	(54.3)	(0.0)	(1.9)	(2528110)
B x C	4	4.2**	2.5*	2.4*	1.2	1.0	0.6	2.3	0.5	0.2	1.8	1.6
		(2.1)	(0.4)	(0.4)	(0.2)	(0.2)	(5730)	(4.3)	(0.8)	(0.0)	(3.1)	(393650)
Subject MS	44	(2.0)	(0.8)	(0.8)	(0.7)	(0.8)	(32770)	(11.0)	(15.8)	(0.1)	(4.7)	(1261710)

*p<.05. **p<.01.

Table 2.9. ANOVA Table for Eye Movement Variables in Block 5

Source	df	F									
		h1	h1t	h2	h2t	mfd	nclust	nfix	leftfix	retfix	rettime
Condition	4	2.9*	4.2**	6.3**	9.1**	2.3	6.1**	19.1**	1.2	2.6*	7.1**
		(0.5)	(0.7)	(1.3)	(1.8)	(36110)	(11.1)	(27.4)	(0.0)	(3.5)	(1902910)
Subject MS	44	(0.5)	(0.5)	(0.4)	(0.5)	(39720)	(5.6)	(11.7)	(0.1)	(4.1)	(908600)

*p<.05. **p<.01.

Table 3.1. ANOVA Table for All Variables in Experiment 4 (n=32)

Source	df	F											
		ratings	h1	h1t	h2	h2t	leftfix	lefttime	mfd	nclust	nfix	retfix	ret
Task (T)	1	82**	0.8	1.4	0.8	1.3	0.2	0.6	0.1	0.6	0.0	0.1	
		(53)	(0.9)	(1.6)	(0.6)	(1.1)	(0.0)	(0.1)	(2558)	(11.9)	(0.2)	(0.7)	(79
Condition (C)	2	733**	21.5**	23.6**	13.3**	15.9**	19.4**	12.8**	4.6**	17.6**	7.6**	2.5	
		(246)	(1.3)	(1.4)	(0.8)	(1.0)	(0.1)	(0.0)	(4410)	(16.0)	(7.1)	(2.4)	(24
Block (B)	2	640**	23.0**	27.3**	22.9**	25.6**	3.7*	4.3*	3.8*	22.5**	16.9**	14.0**	7.
		(340)	(2.5)	(3.1)	(1.9)	(2.4)	(0.0)	(0.1)	(8310)	(38.2)	(22.7)	(21.4)**	(117
C x T	2	15.1**	5.6**	8.2**	4.7**	6.4**	0.4	0.3	4.6**	7.1**	5.6**	4.0*	1
		(5.1)	(0.3)	(0.5)	(0.3)	(0.4)	(0.0)	(0.0)	(4400)	(6.4)	(5.2)	(3.8)	(17
B x T	2	10.7**	5.2**	6.1**	11.7**	13.7**	0.3	0.3	0.4	7.7**	13.1**	0.0	4
		(5.7)	(0.6)	(0.7)	(1.0)	(1.3)	(0.0)	(0.0)	(927)	(13.0)	(17.5)	(0.1)	(67
C x B	4	288**	2.2+	2.4*	3.0*	4.1**	13.2**	13.1**	3.9**	3.1**	5.1**	0.6	1
		(67)	(0.1)	(0.1)	(0.1)	(0.1)	(0.0)	(0.0)	(1590)	(1.7)	(2.4)	(0.5)	(11
C x B x T	4	21**	2.1	2.8*	4.4**	5.4**	1.7	3.2*	3.7**	3.7**	4.3**	1.0	1
		(4.9)	(0.1)	(0.1)	(0.1)	(0.2)	(0.0)	(0.0)	(1480)	(1.9)	(2.0)	(0.7)	(16

+p<0.08. *p<.05. **p<.01.

Table 3.2. ANOVA Table for All Variables, Scene Task Group (n=16)

Source	df	F											
		ratings	h1	h1t	h2	h2t	leftfix	lefttime	mfd	nfix	nclust	retfix	rettime
Condition (C)	2	302**	19.1**	24.1**	11.4**	13.9**	8.1**	5.5**	5.5**	9.6**	17.3**	3.8*	0.5
		(132)	(1.4)	(1.8)	(1.0)	(1.3)	(0.0)	(0.0)	(8210)	(11.8)	(21.4)	(5.3)	(73450)
Block (B)	2	428**	17.7**	21.5**	21.1**	24.1**	2.6	2.8+	2.1	20.2**	20.2**	4.5*	3.0+
		(203)	(2.7)	(3.4)	(2.7)	(3.4)	(0.0)	(0.0)	(7390)	(39.4)	(47.6)	(10.2)	(732250
C x B	4	136**	4.7**	5.7**	6.2**	7.7**	6.3**	9.6**	4.7**	6.5**	7.0**	0.6	2.1
		(36)	(0.1)	(0.2)	(0.2)	(0.3)	(0.0)	(0.0)	(2720)	(4.0)	(3.3)	(0.4)	(177540
Subject MS	15	(0.7)	(0.5)	(0.6)	(0.4)	(0.7)	(0.1)	(0.1)	(10160)	(17.6)	(10.1)	(3.0)	(104646

+p<0.08. *p<.05. **p<.01.

Table 3.3. F Values (MS) from Contrasts on All Variables between Conditions within Block 3, Scene Recognition Task (n=16)

Conditions	F												
	ratings	h1	h1t	h2	h2t	leftfix	lefttime	mfd	nfix	nclust	retfix	rettime	s1
Same v.	661**	45.2**	60.0**	55.7**	67.6**	0.1	0.1	33.8**	48.0**	59.3**	1.5	4.1*	10.8**
Control	(174)	(1.4)	(2.0)	(1.7)	(2.3)	(0.0)	(0.0)	(19530)	(29.4)	(27.8)	(1.1)	(356963)	(0.0)
Same v.	0.3	2.7	4.1*	4.5*	4.9*	2.3	3.2+	3.1+	12.2**	5.8*	0.7	7.1**	1.0
Translate	(0.1)	(0.1)	(0.1)	(0.1)	(0.2)	(0.0)	(0.0)	(1810)	(7.5)	(2.7)	(0.5)	(609600)	(0.0)
Translate v.	633**	25.9**	32.8**	28.6**	36.2**	3.1+	4.6*	16.4**	11.8**	27.9**	4.3*	0.4	18.4**
Control	(166)	(0.8)	(1.1)	(0.9)	(1.2)	(0.0)	(0.0)	(9450)	(7.3)	(13.1)	(3.1)	(33600)	(0.0)

+p<.08. *p<.05. **p<.01.

Table 3.4. F Values (MS) from Contrasts between Conditions for Eye Movement Variables, Scene Recognition Task (n=16)

Source	F					
	h1	h1t	h2t	mfd	nfix	nclust
Same v. T1	NS	NS	NS	NS	NS	NS
Same v. T2	5.0 (0.6)*	7.1 (0.8)**	7.8 (0.9)**	NS	4.9 (11.6)*	5.1 (9.7)*
T1 v. T2	3.8 (0.5)+	5.1 (0.6)*	6.4 (0.8)*	NS	NS	NS
T2 v. New	NS	NS	NS	NS	NS	5.6 (3.1)*

NS=Not significant

+p<.08. *p<.05. **p<.01.

Table 3.5 ANOVA Table for Selected Variables, Picture Task Group (n=16)

Source	df	F				
		ratings	h2t	mfd	nfix	nclust
Condition (C)	2	508**	2.6+	1.5	1.1	1.9
		(119)	(0.1)	(598)	(0.7)	(1.1)
Block (B)	2	243**	4.6*	2.1	0.1	3.5*
		(144)	(0.2)	(1842)	(0.3)	(3.6)
C x B	4	179**	0.6	1.5	0.7	0.5
		(36)	(0.0)	(353)	(0.3)	(0.3)
Subject MS	15	(0.6)	(1.1)	(31970)	(45.4)	(32.8)

+p<0.08. *p<.05. **p<.01.

Table 4.1. ANOVA Table for Selected Variables in Experiment 4 (n=24)

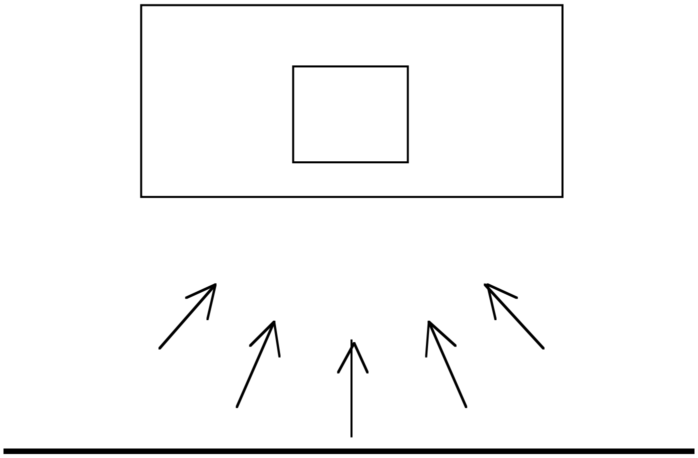
Source	df	F								
		ratings	nfix	h2t	nclust	mfd	leftfix	lefttime	retfix	rettime
Condition (C)	2	342**	5.3**	6.0**	4.9**	0.9	1.8	3.0+	2.1	1.9
		(188)	(2.5)	(0.3)	(4.8)	(263)	(0.0)	(0.0)	(1.6)	(19910)
Block (B)	2	578**	4.2*	14.1**	11.5**	2.6+	1.4	1.6	2.9+	1.0
		(307)	(6.3)	(0.9)	(23.8)	(2320)	(0.0)	(0.0)	(5.6)	(246990)
C x B	4	135**	3.1*	2.4+	2.1+	0.9	2.2+	1.6	1.5	1.6
		(48)	(1.5)	(0.1)	(1.7)	(196)	(0.0)	(0.0)	(1.0)	(145530)
Subject MS	23	(1.4)	(13.7)	(0.3)	(12.1)	(15230)	(0.1)	(0.1)	(5.3)	(730436)

+p<.08. *p<.05. **p<.01.

Table 4.2. F Values (MS) from Contrasts on Selected Variables between Conditions within Block 3 (n=24)

Conditions	F			
	ratings	nfix	h2t	nclust
Same v. Control	728 (256)**	20.0 (9.4)**	17.1 (0.8)**	10.8 (8.6)**
Same v. Reflect	0.3 (0.1)	5.7 (2.7)*	4.3 (0.2)+	0.5 (0.4)
Reflect v. Control	699 (246)**	4.4 (2.1)*	4.2 (0.2)*	6.8 (5.4)**

+p<.08. *p<.05. **p<.01.



Note: All five images have the same focal point and subtend the same visual angle

Figure 1.1. Schematic Depicting Viewer Perspectives and Direction of Camera at Each Viewpoint

Figure 1.2. Examples of Stimuli Used in Experiment 1

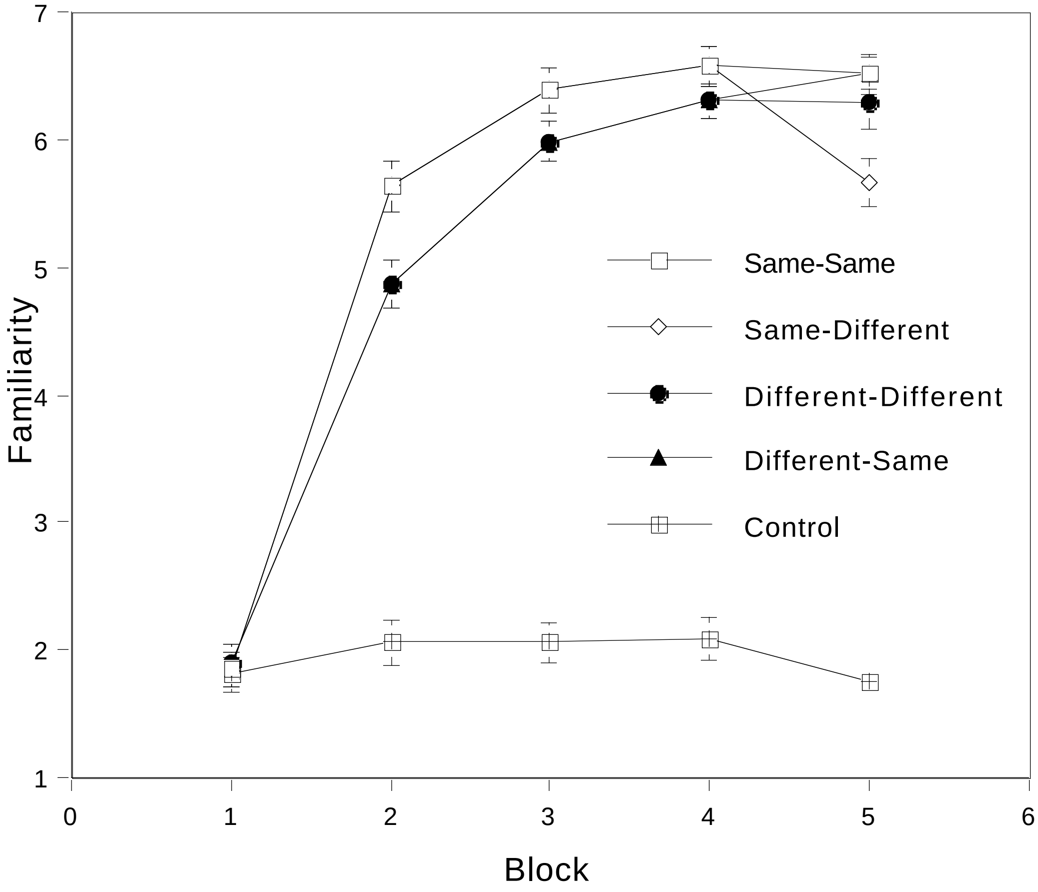


Figure 1.3. The Effect of Viewpoint and Exposure on Familiarity Ratings of Naturalistic Scenes

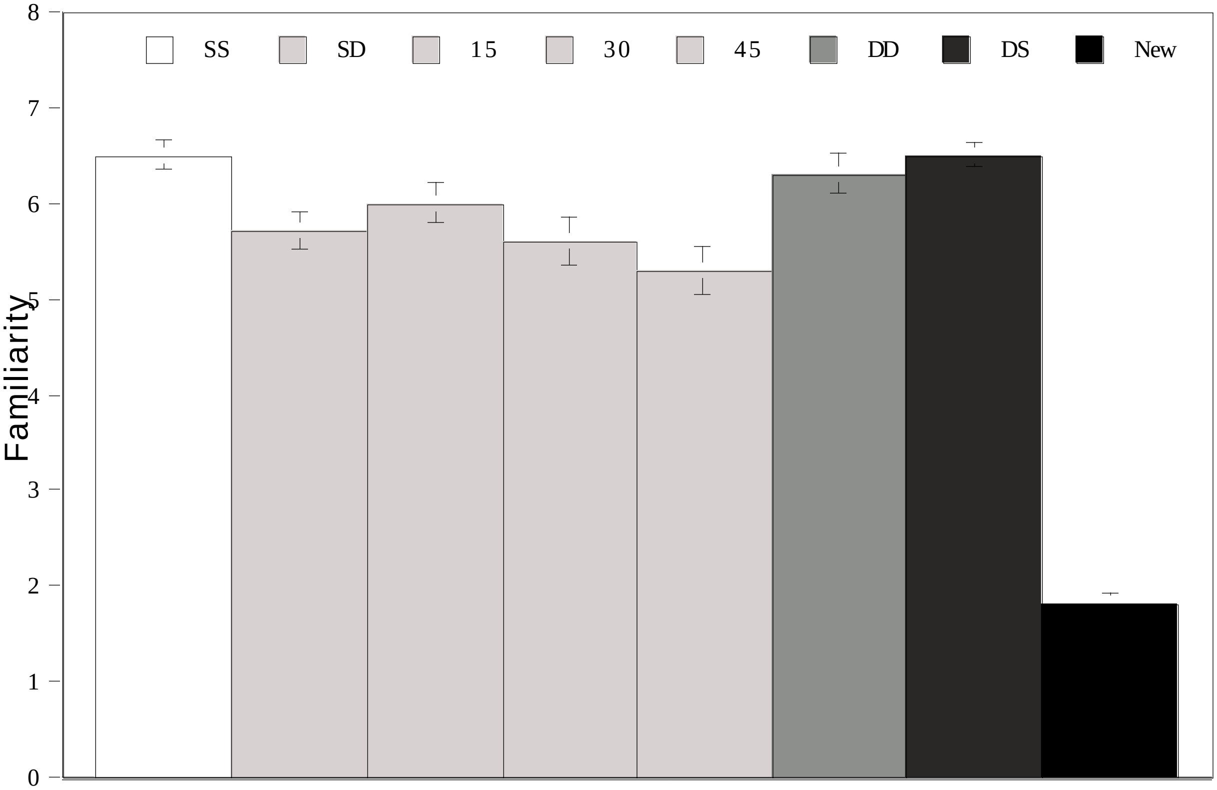


Figure 1.4. Familiarity Ratings in Block 5, with Levels of Rotation in the Same-Different (SD) Condition

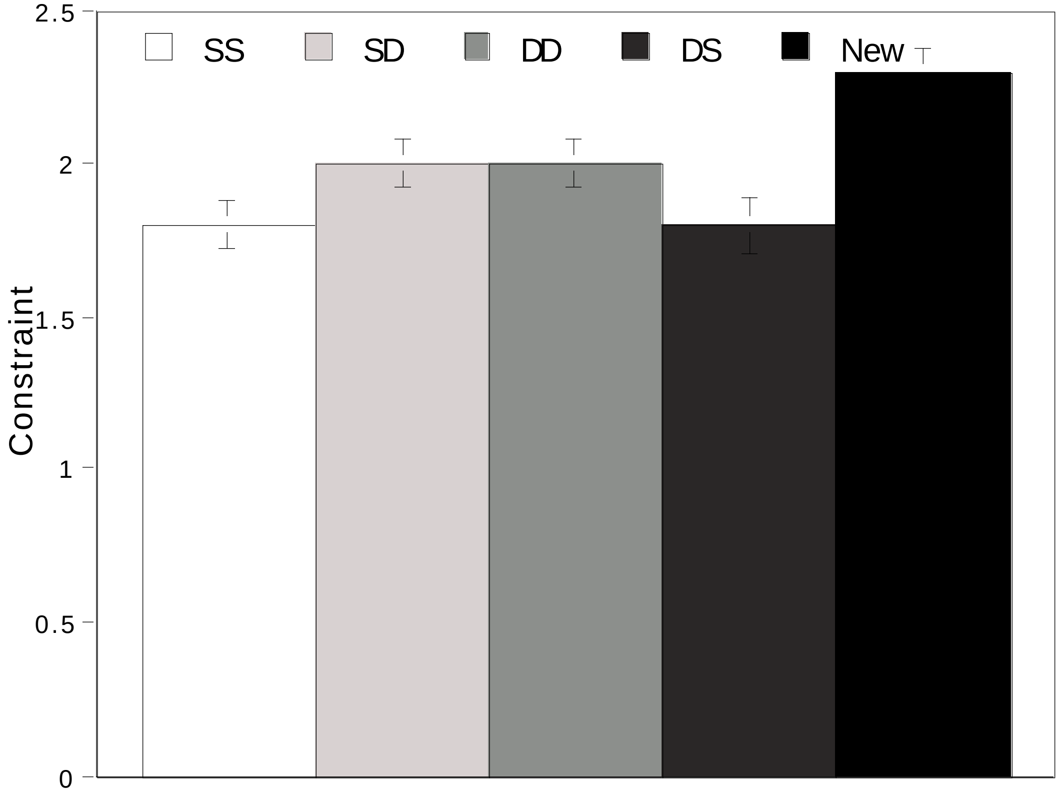


Figure 1.5. H1 Values in Block 5

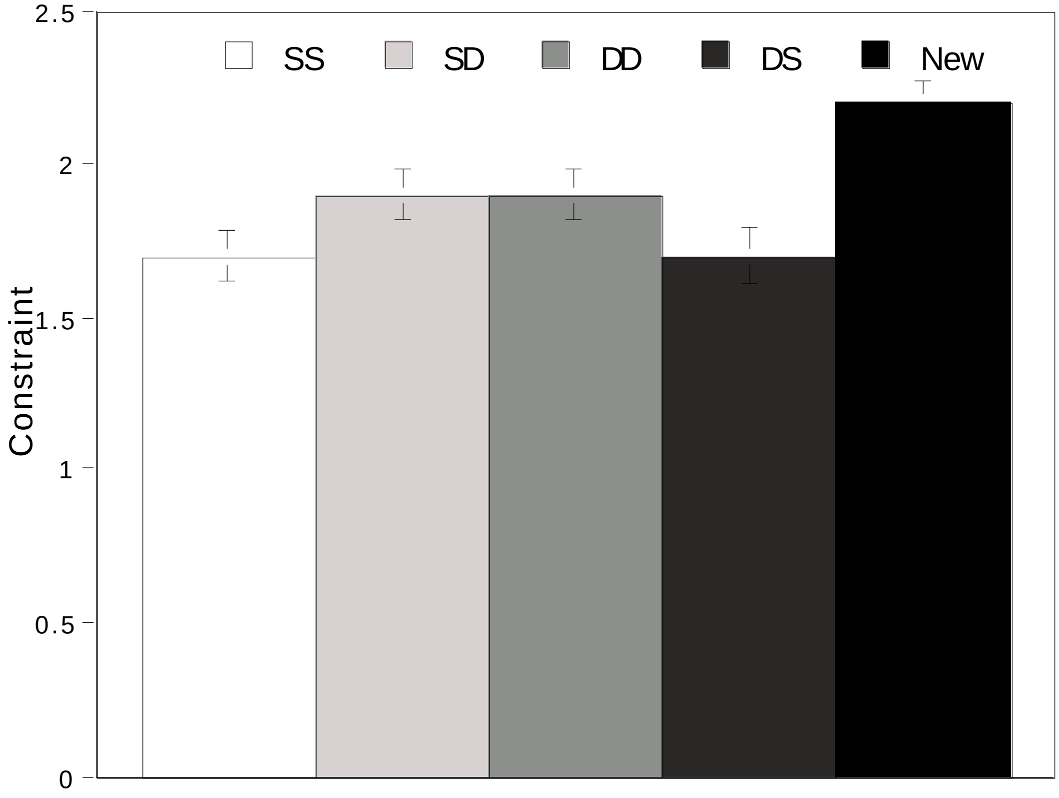


Figure 1.6. H1t Values in Block 5

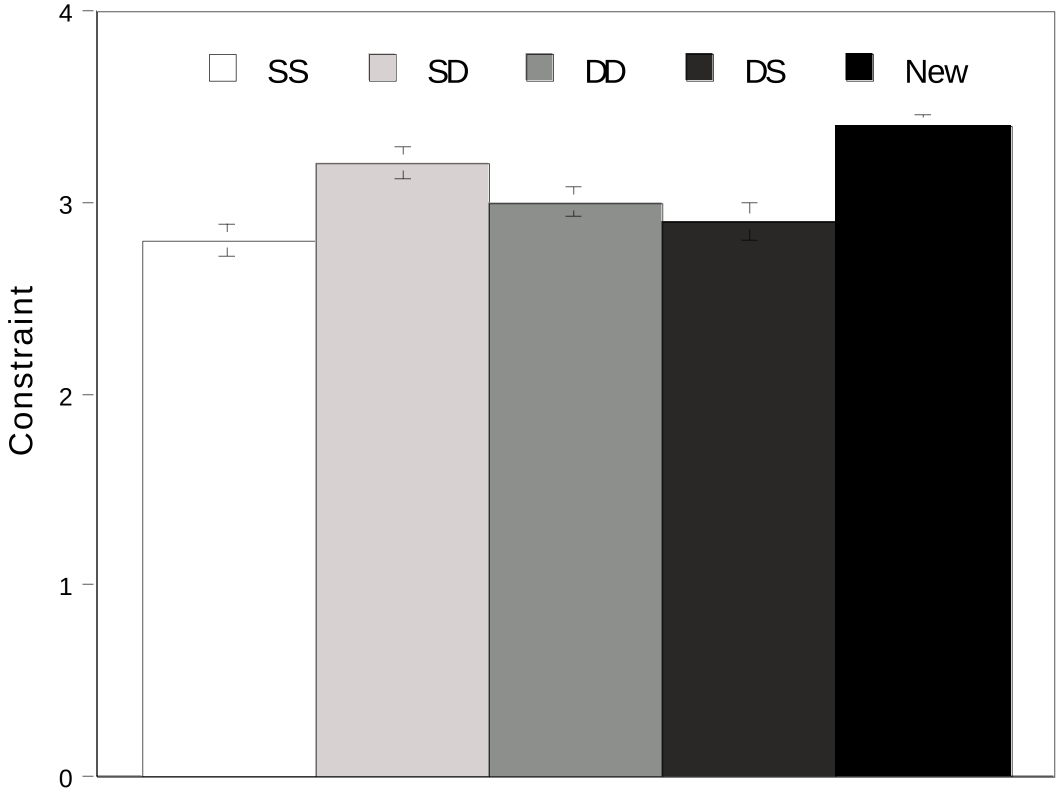


Figure 1.7. H2 Values in Block 5

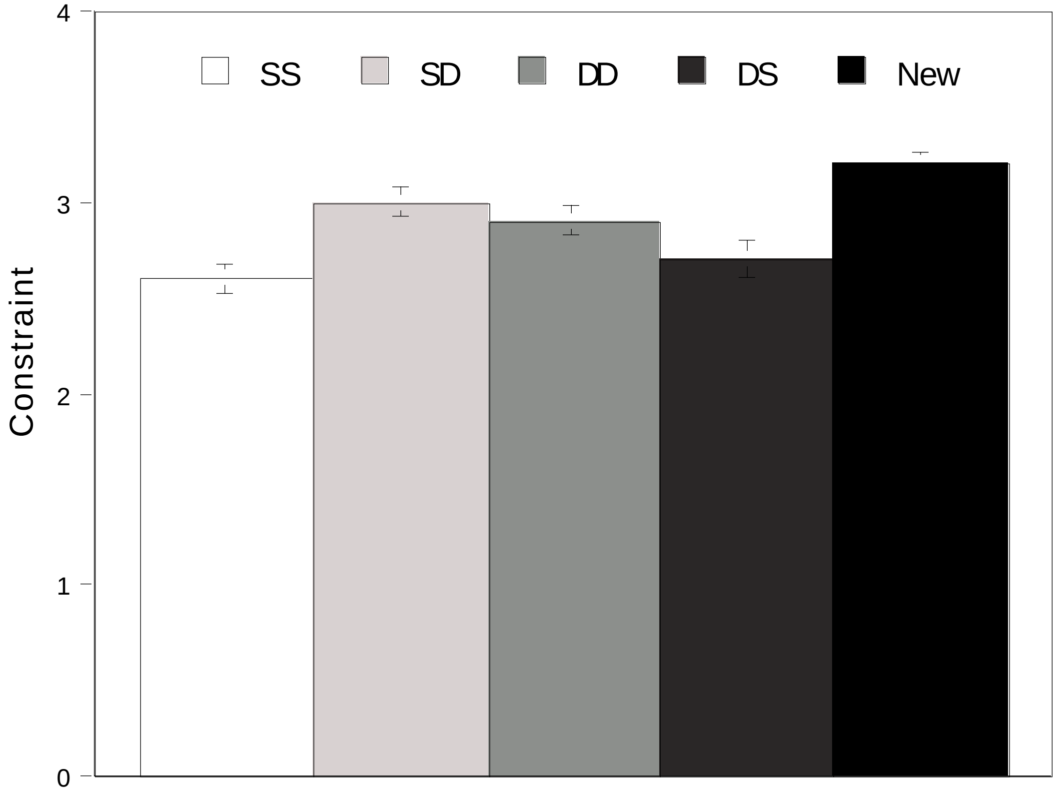


Figure 1.8. H2t Values in Block 5

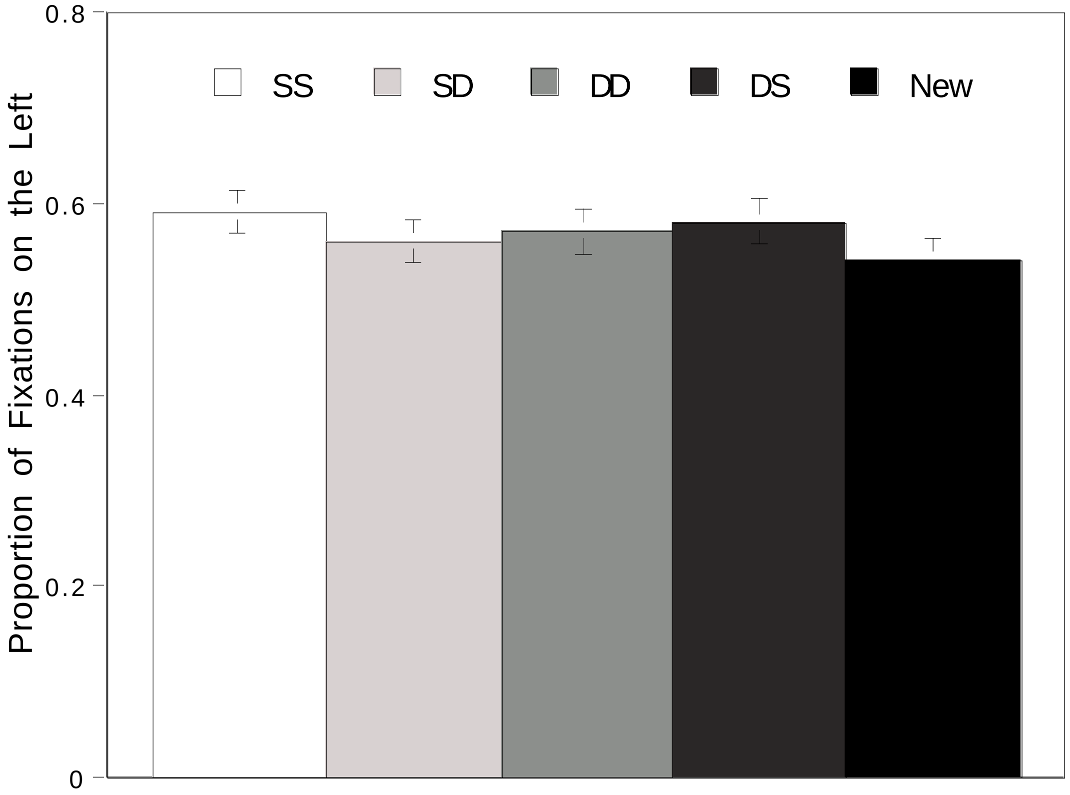


Figure 1.9. Proportion of Fixations on the Left Side of the Picture in Block 5

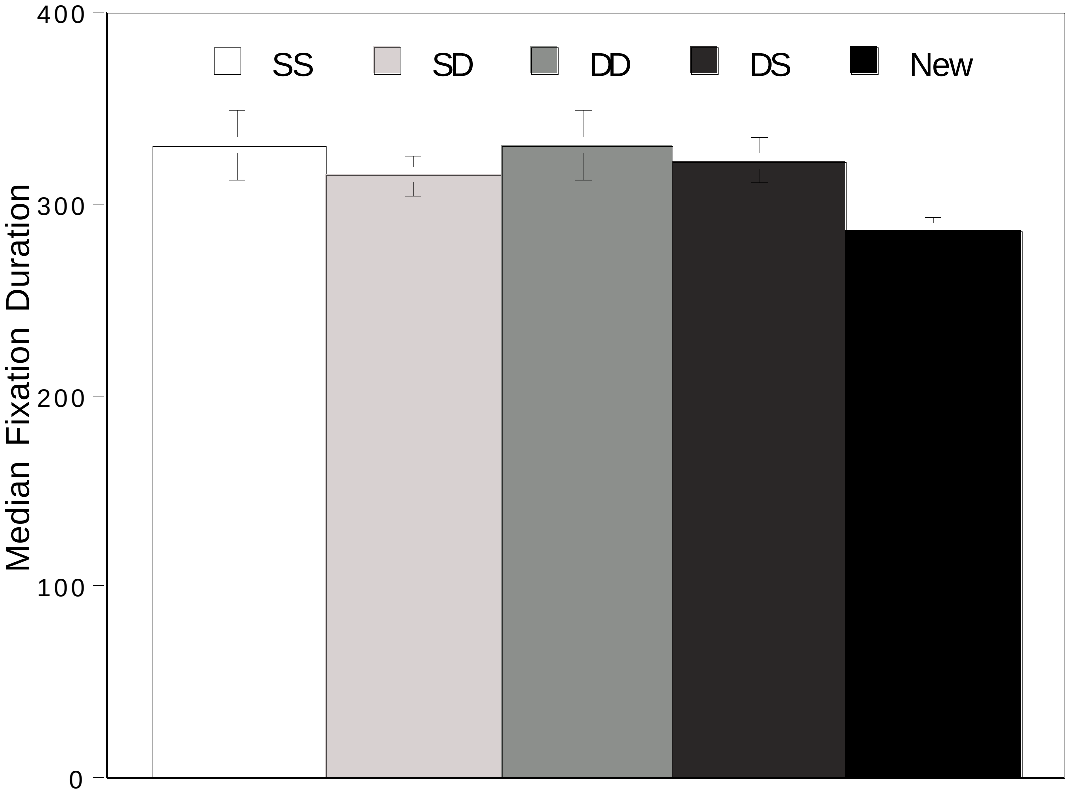


Figure 1.10. Median Fixation Duration in Block 5

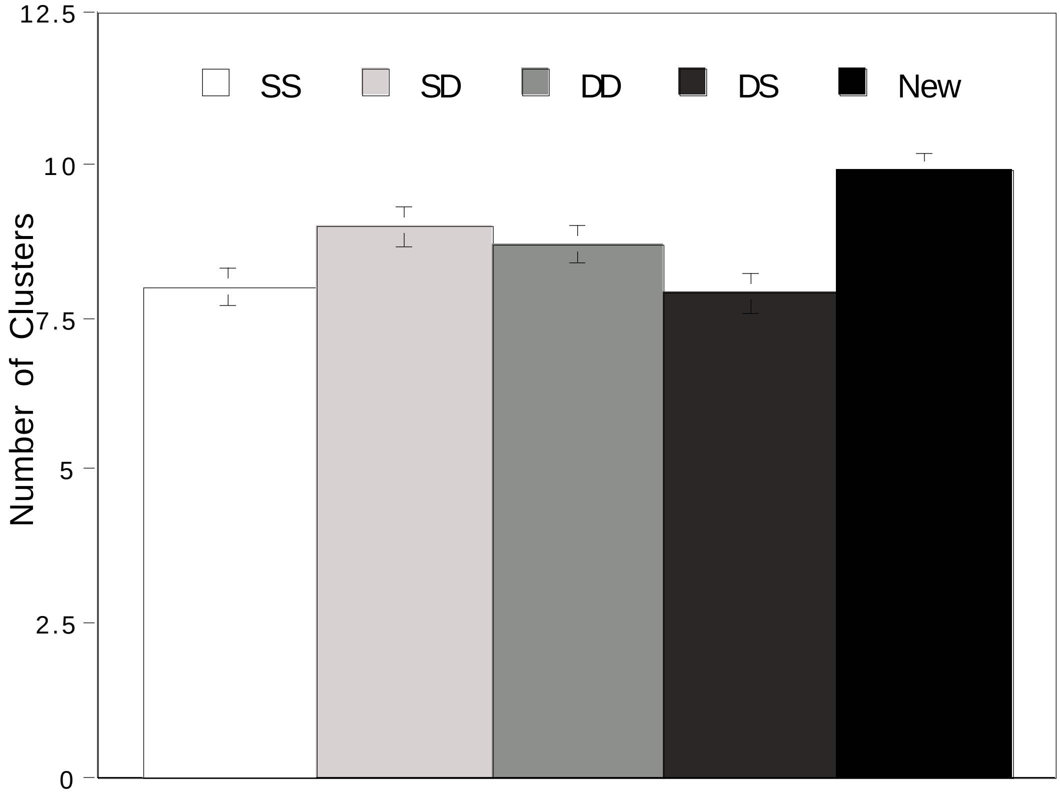


Figure 1.11. Number of Clusters in Block 5

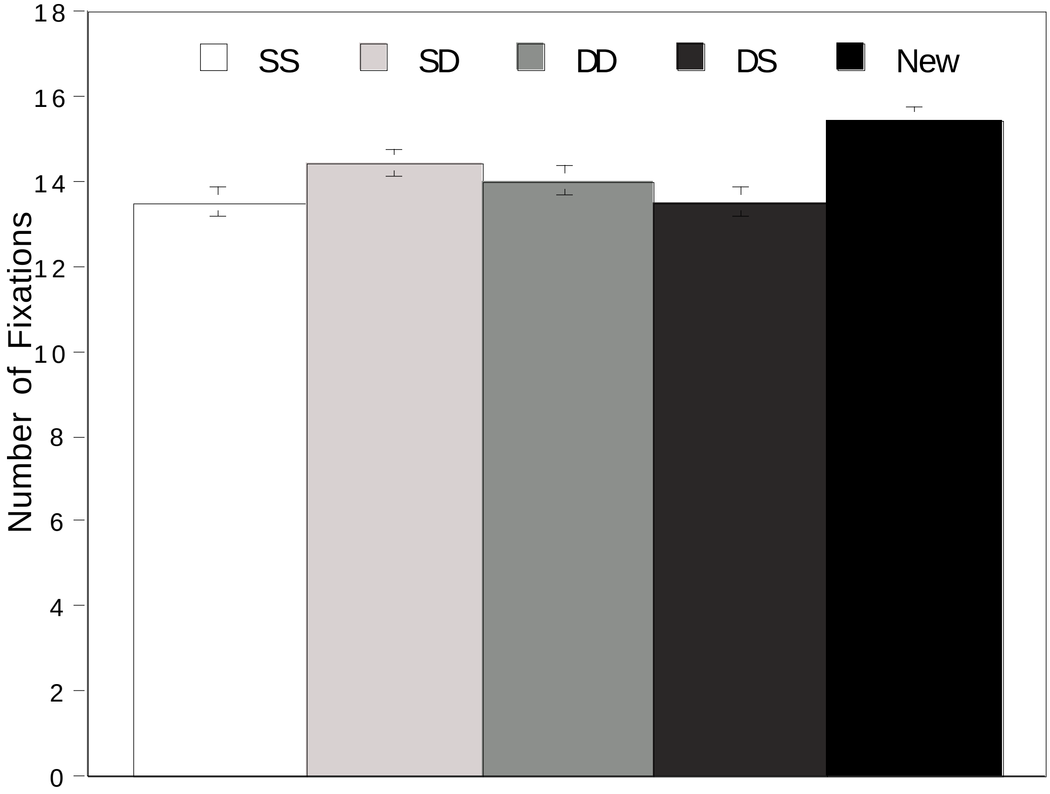


Figure 1.12. Number of Fixations in Block 5

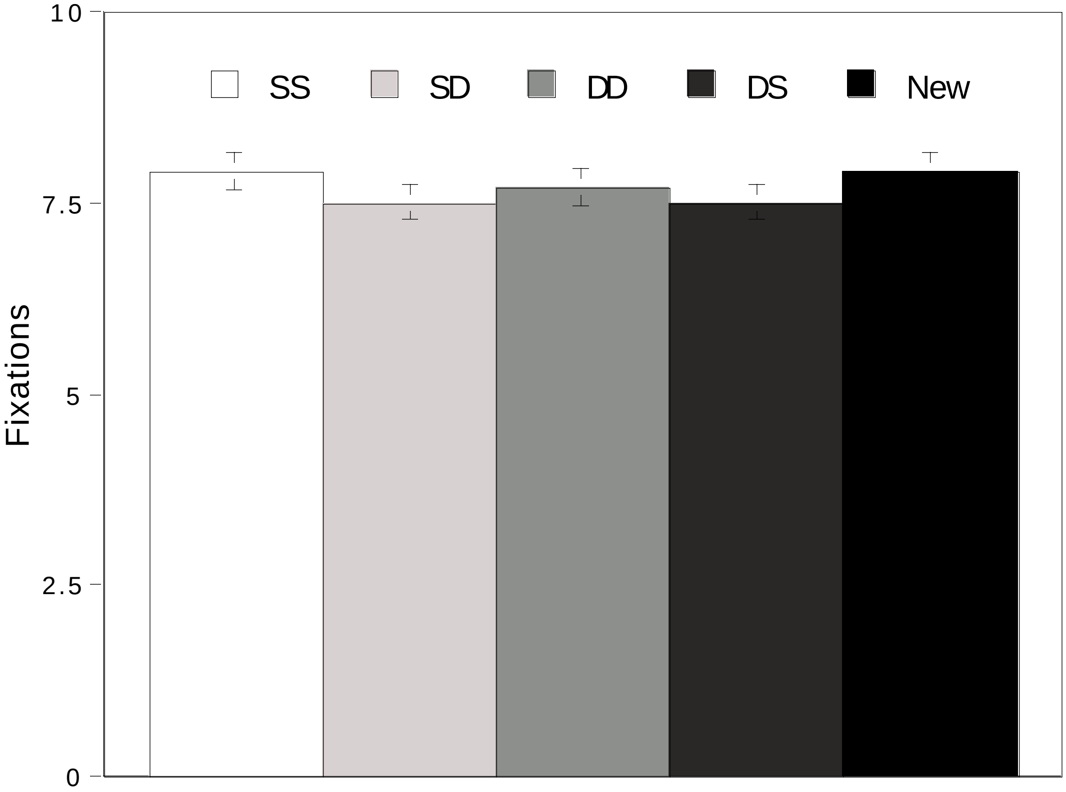


Figure 1.13. Number of Fixations to Return to Original Fixation Location in Block 5

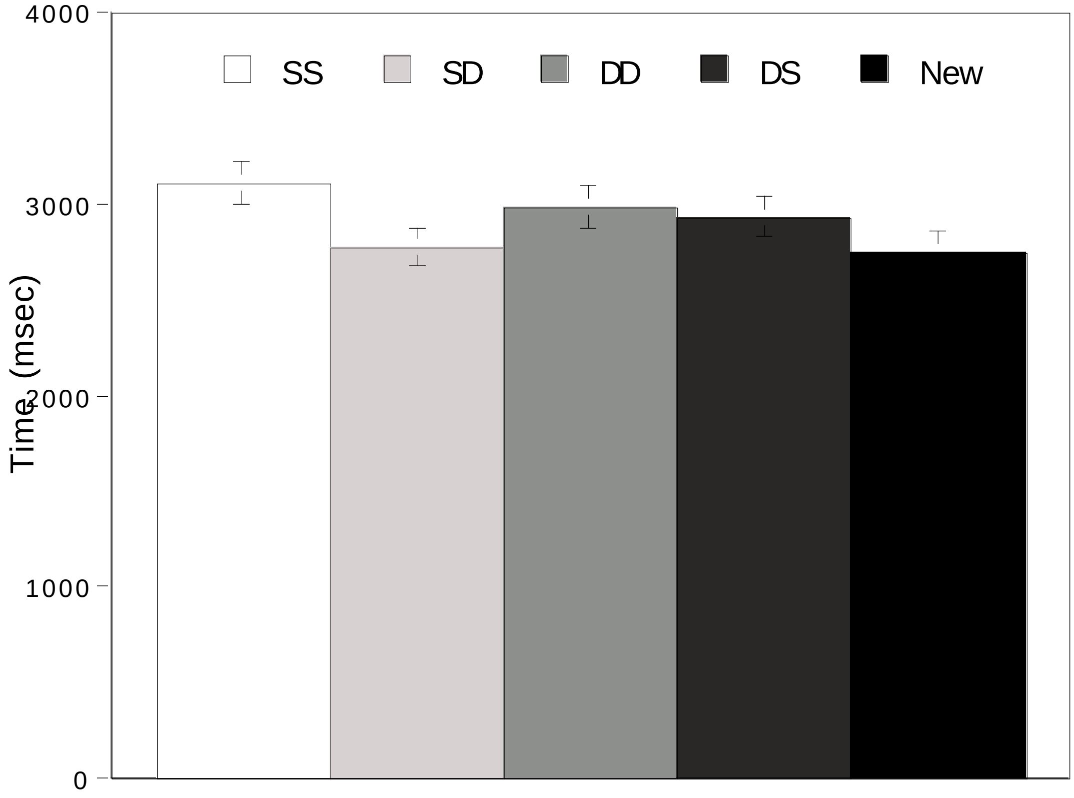


Figure 1.14. Time to Return to Original Fixation Location in Block 5

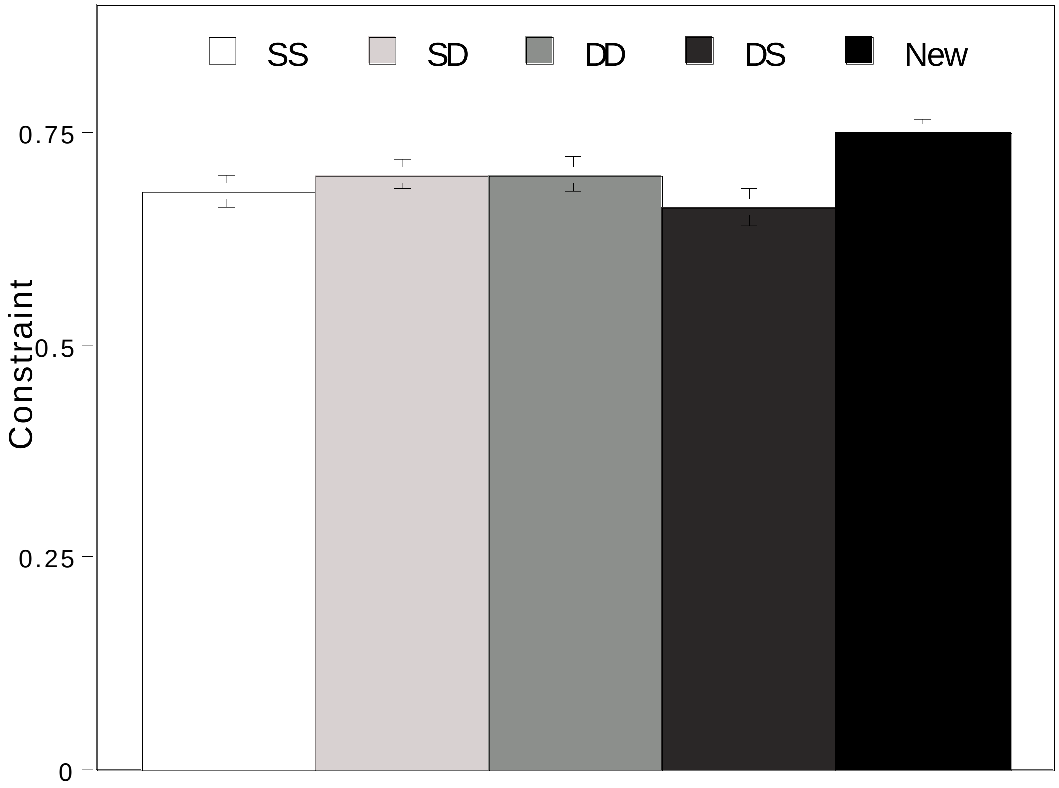


Figure 1.15. S1 Values in Block 5

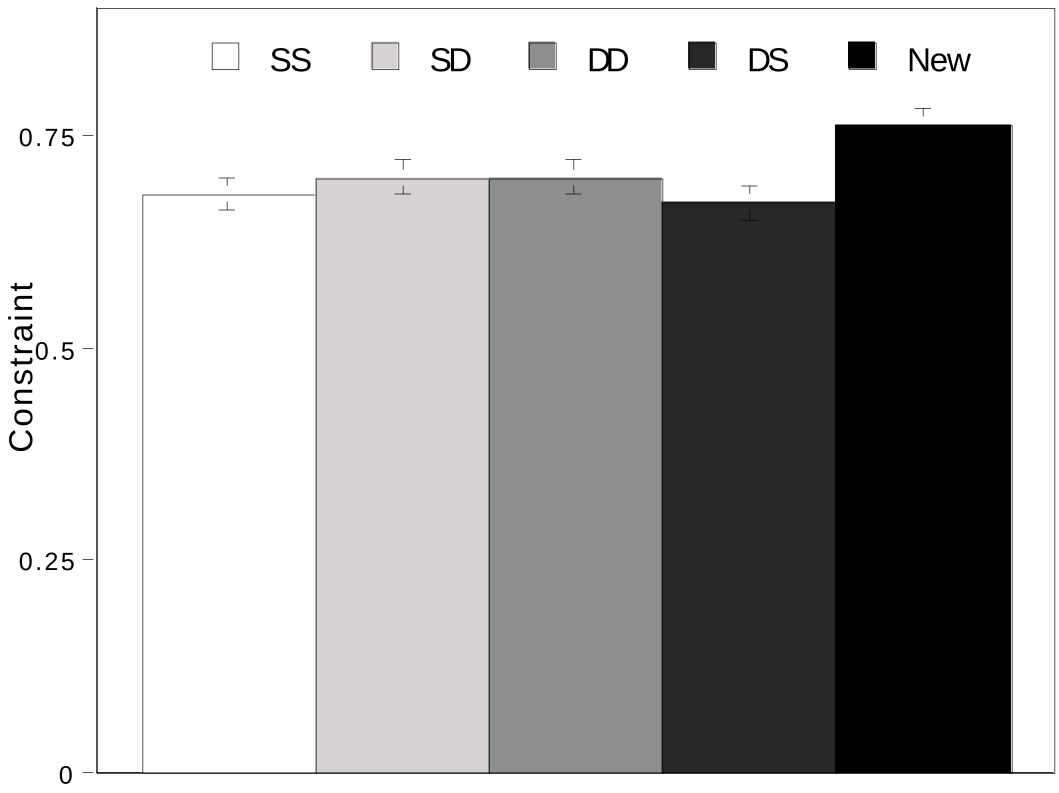


Figure 1.16. S1t Values in Block 5

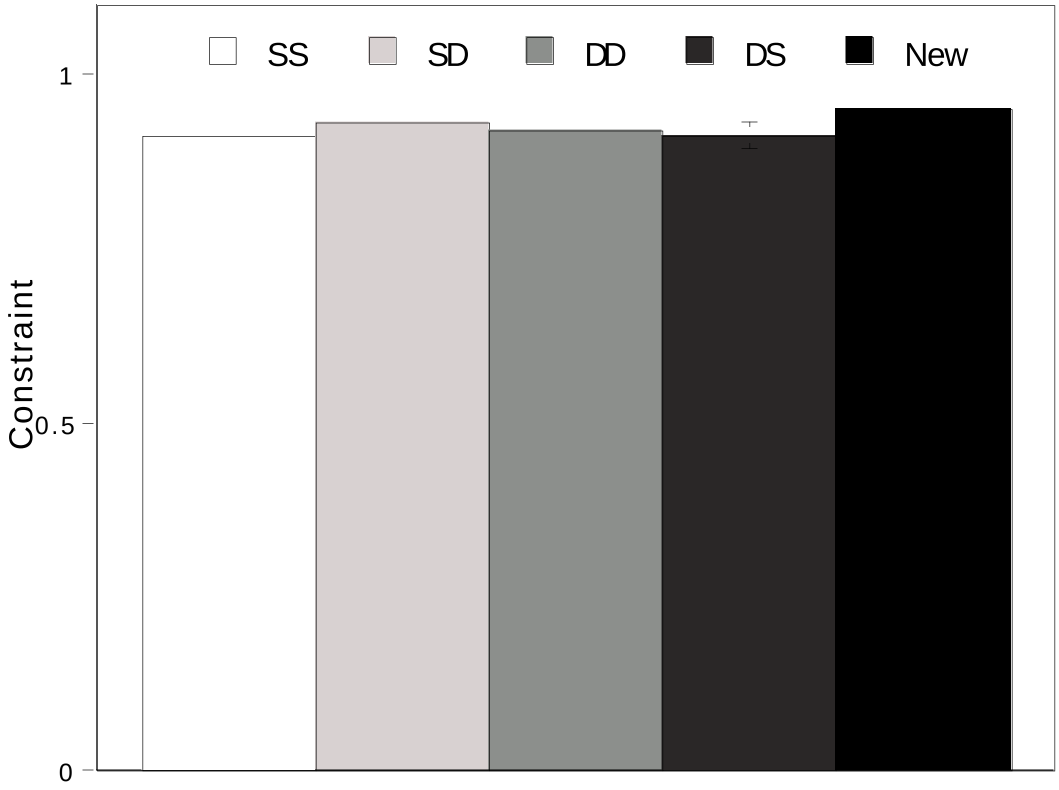


Figure 1.17. S2 Values in Block 5

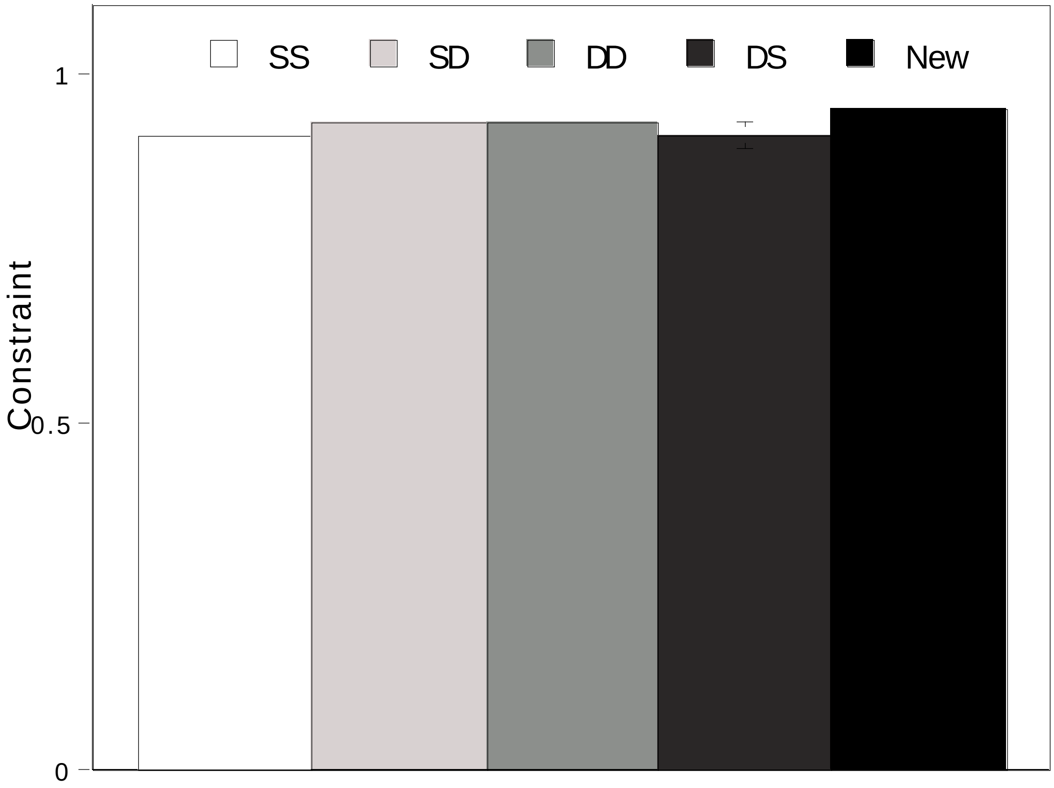


Figure 1.18. S2t Values in Block 5

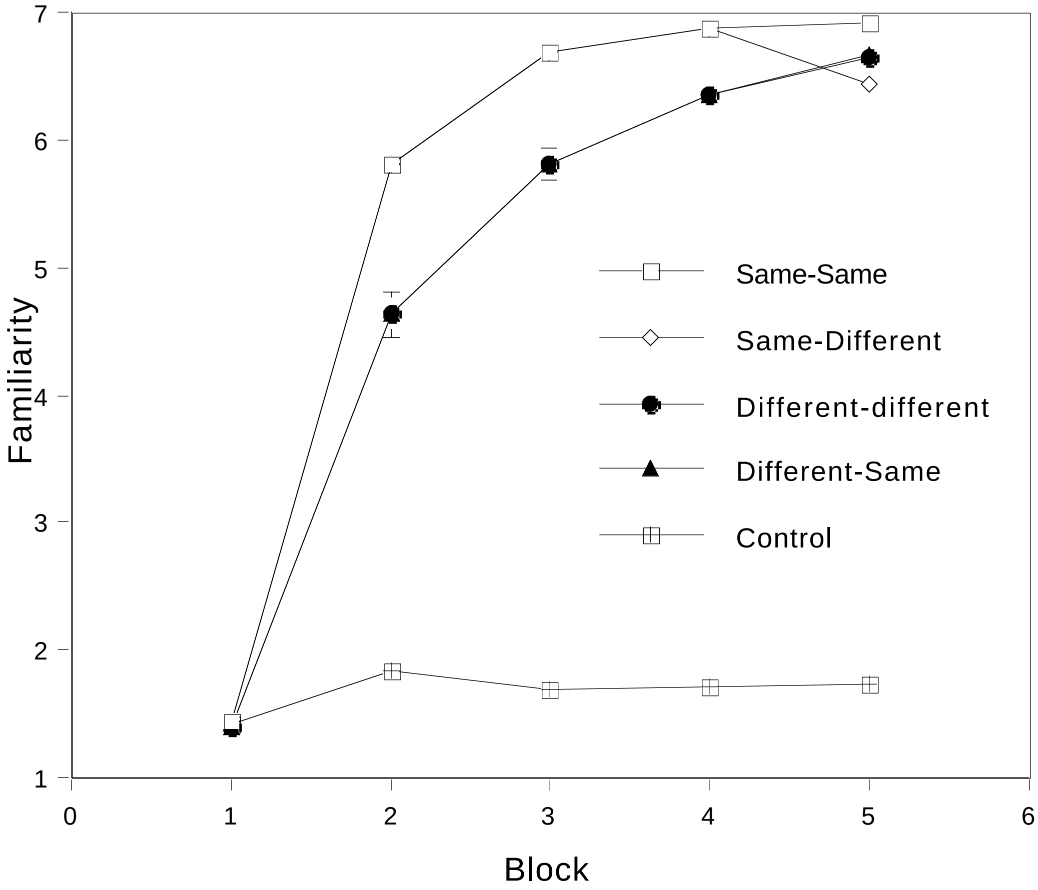


Figure 2.1. The Effect of Viewpoint and Exposure on Familiarity Ratings of Scenes with Backgrounds Removed

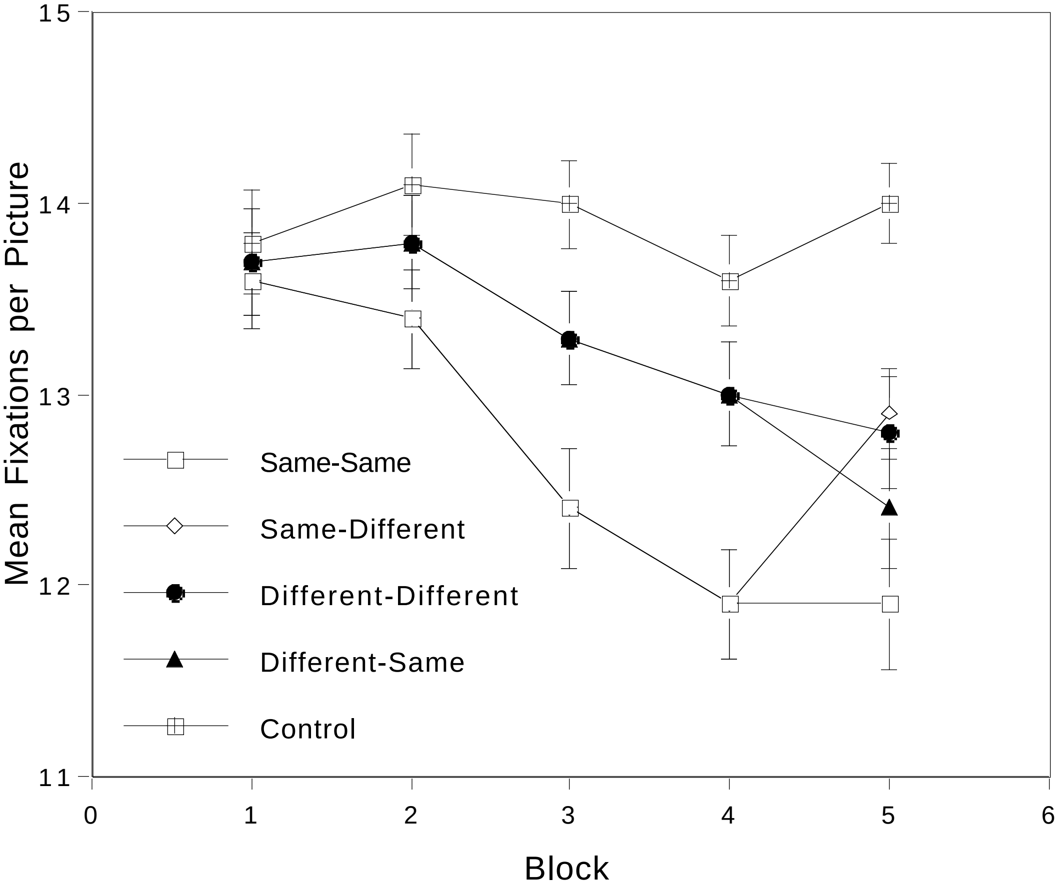


Figure 2.2. The Effect of Viewpoint and Exposure on the Mean Number of Fixations per Picture

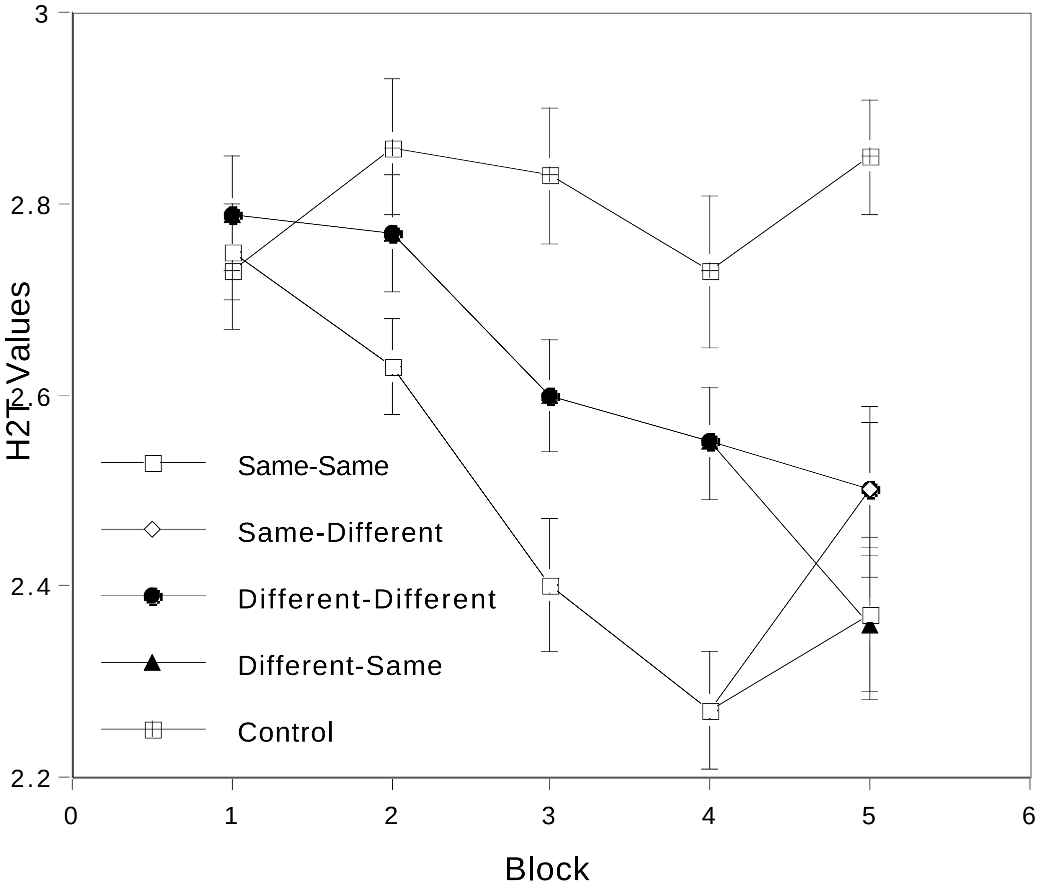


Figure 2.3. The Effect of View Condition on the H2T Variable Across Five Blocks of Exposure

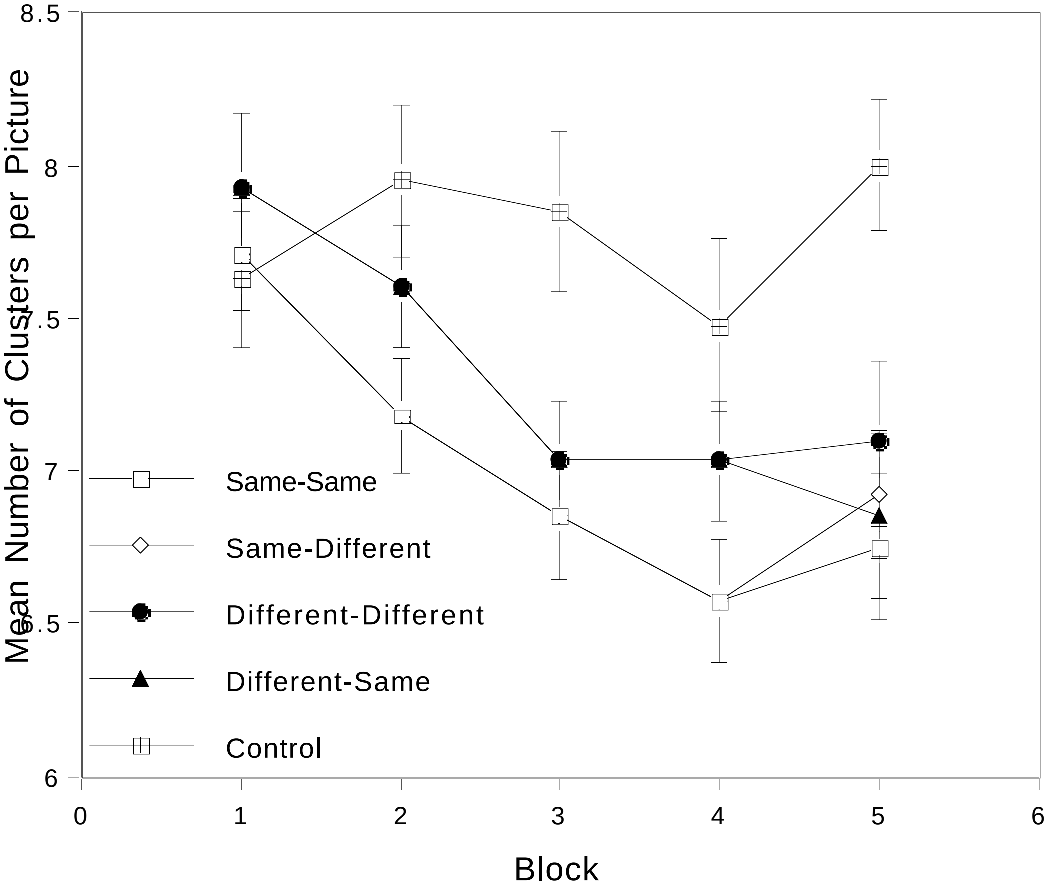


Figure 2.4. The Mean Number of Fixation Clusters Across Five Blocks by View Condition

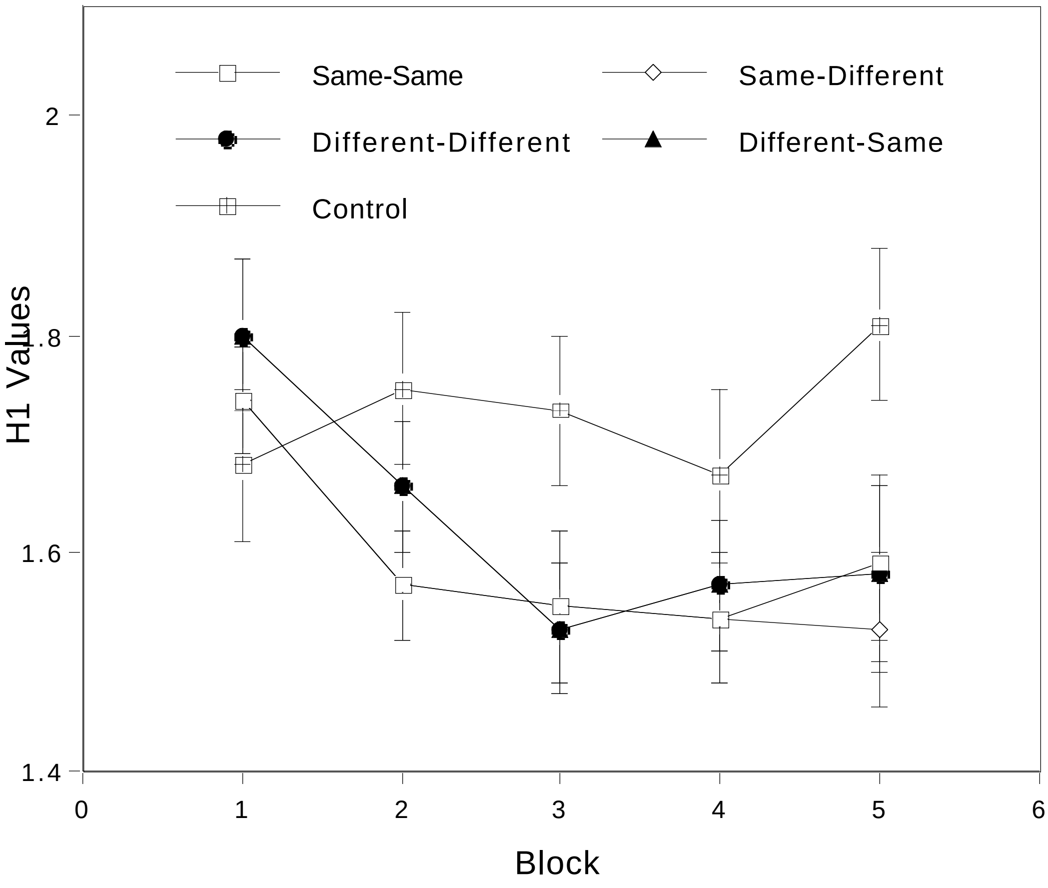


Figure 2.5. The Effect of Viewpoint and Exposure on the H1 Entropy Measure

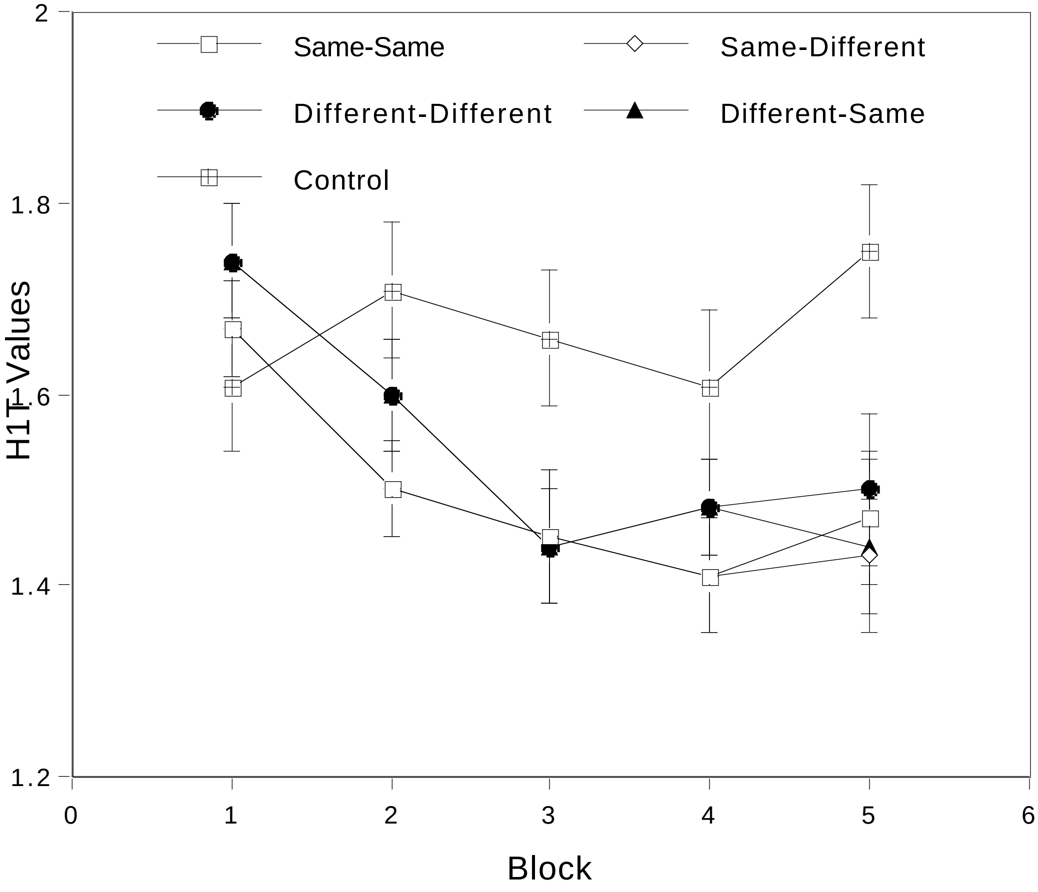


Figure 2.6. The Effect of Viewpoint and Exposure on the H1T Entropy Measure

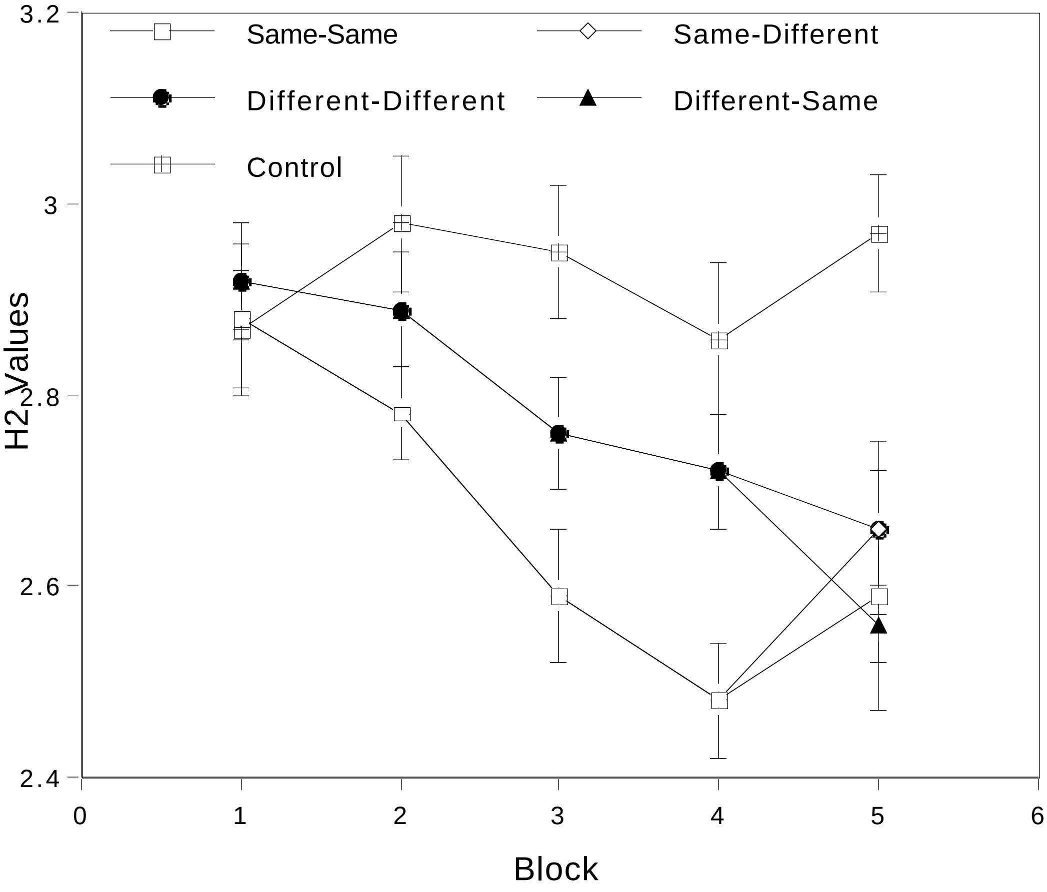


Figure 2.7. The Effect of Viewpoint and Exposure on the H2 Entropy Measure

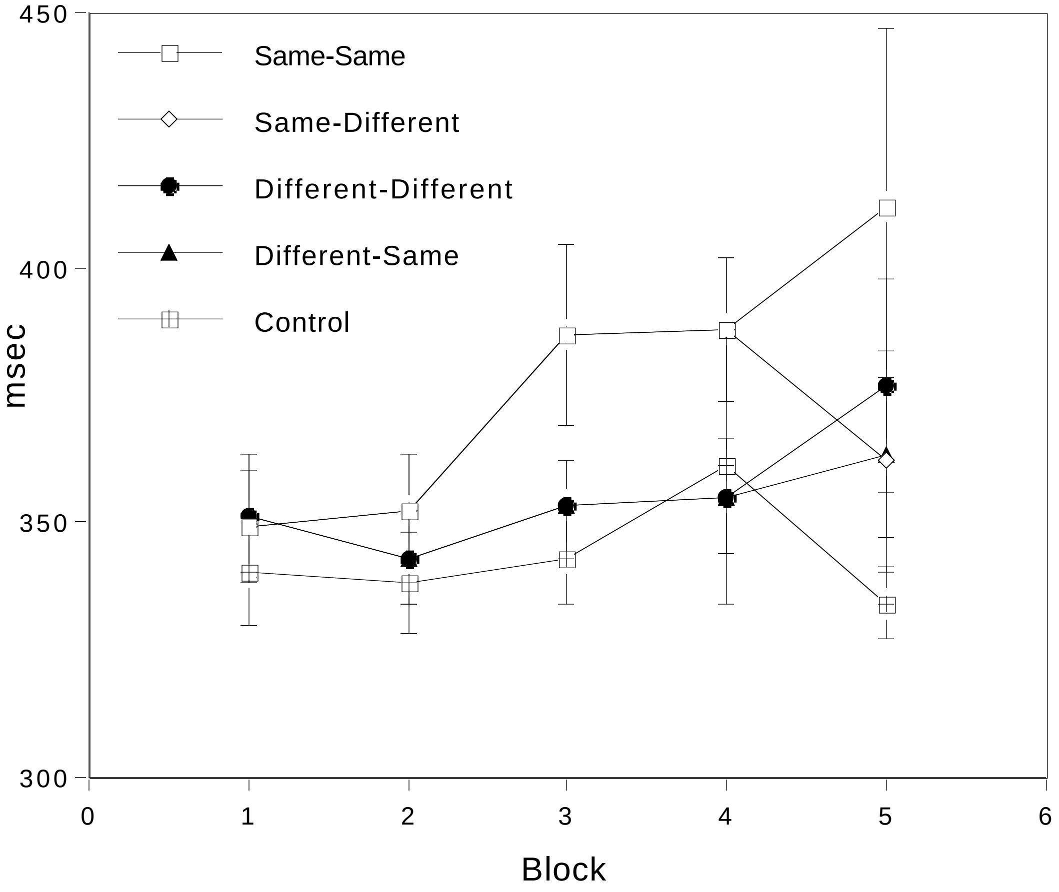
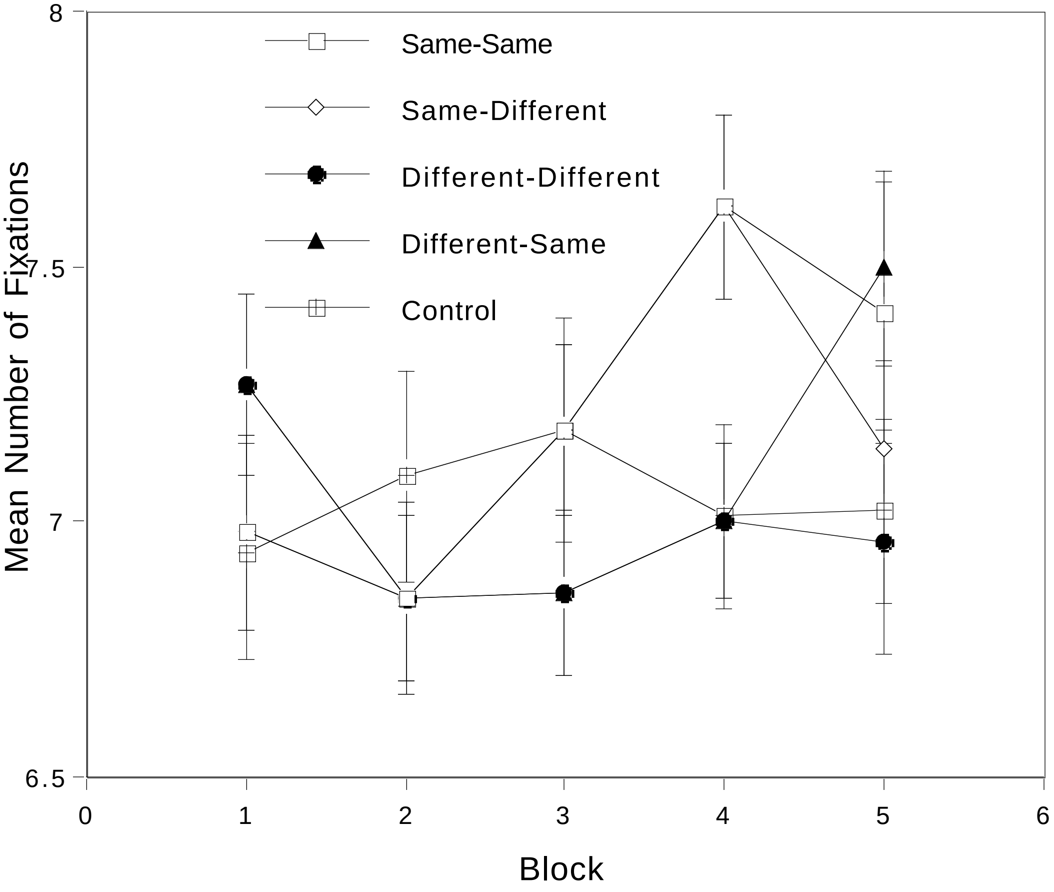


Figure 2.8. The Effect of Viewpoint and Exposure on Median Fixation Duration



<u>Figure 2.9. The Effect of Viewpoint and Exposure on the Number of Fixations it Takes to Return to the Original Fixation Location</u>

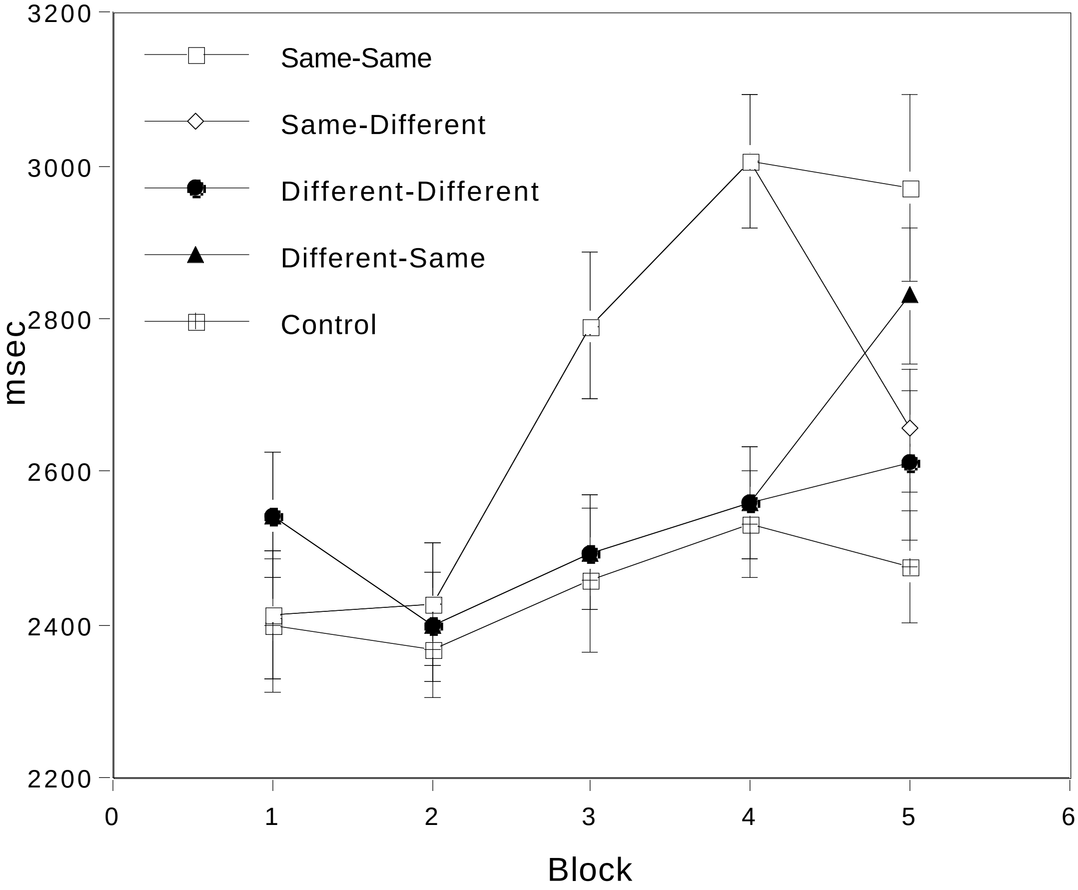


Figure 2.10. The Effect of Viewpoint and Exposure on the Amount of Time it Takes to Return to the Original Fixaiton Location

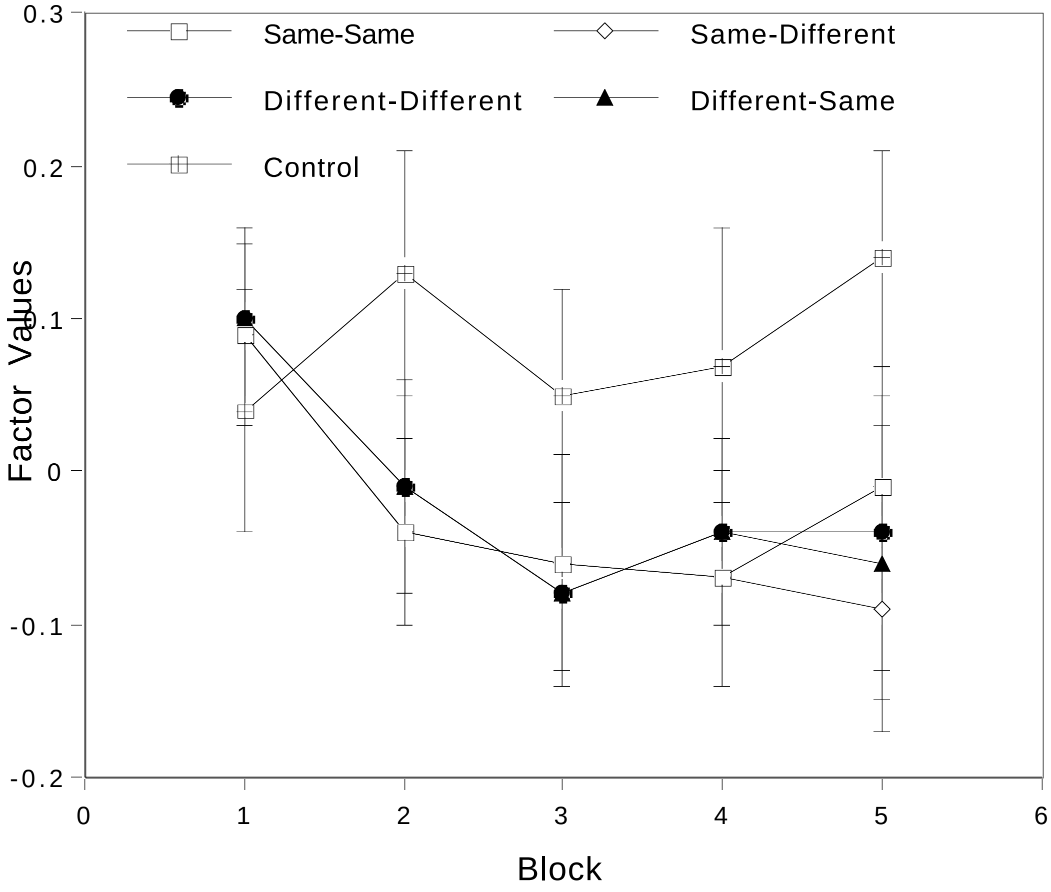


Figure 2.11. The Effect of Viewpoint and Exposure on the Path Factor

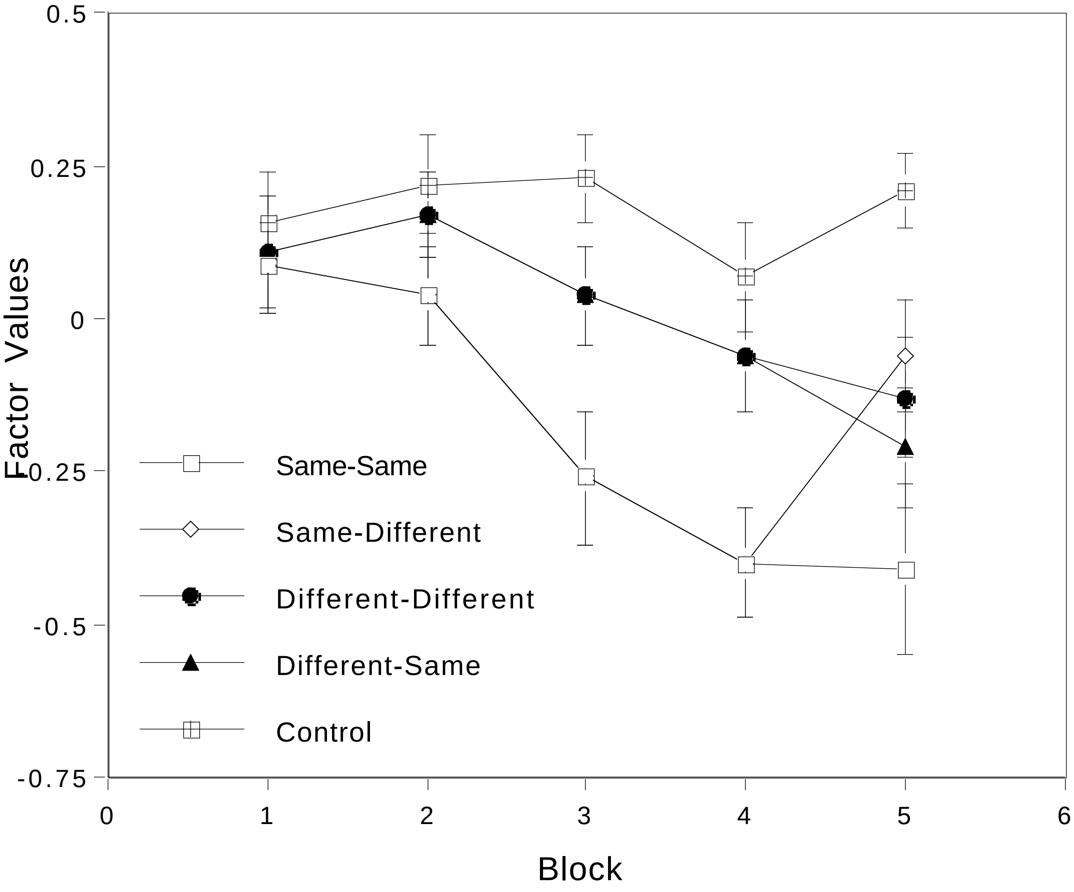


Figure 2.12. The Effect of Viewpoint and Exposure on the Fixation Factor

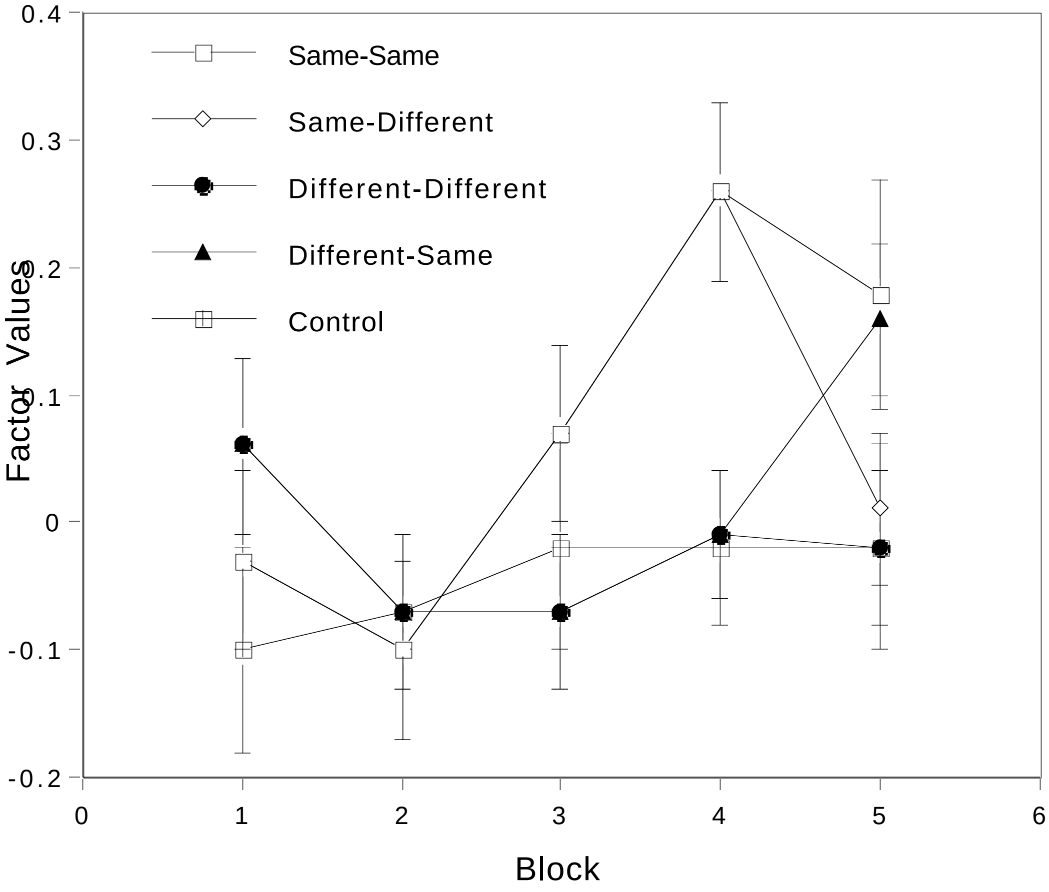


Figure 2.13. The Effect of Viewpoint and Exposure on the Return Factor

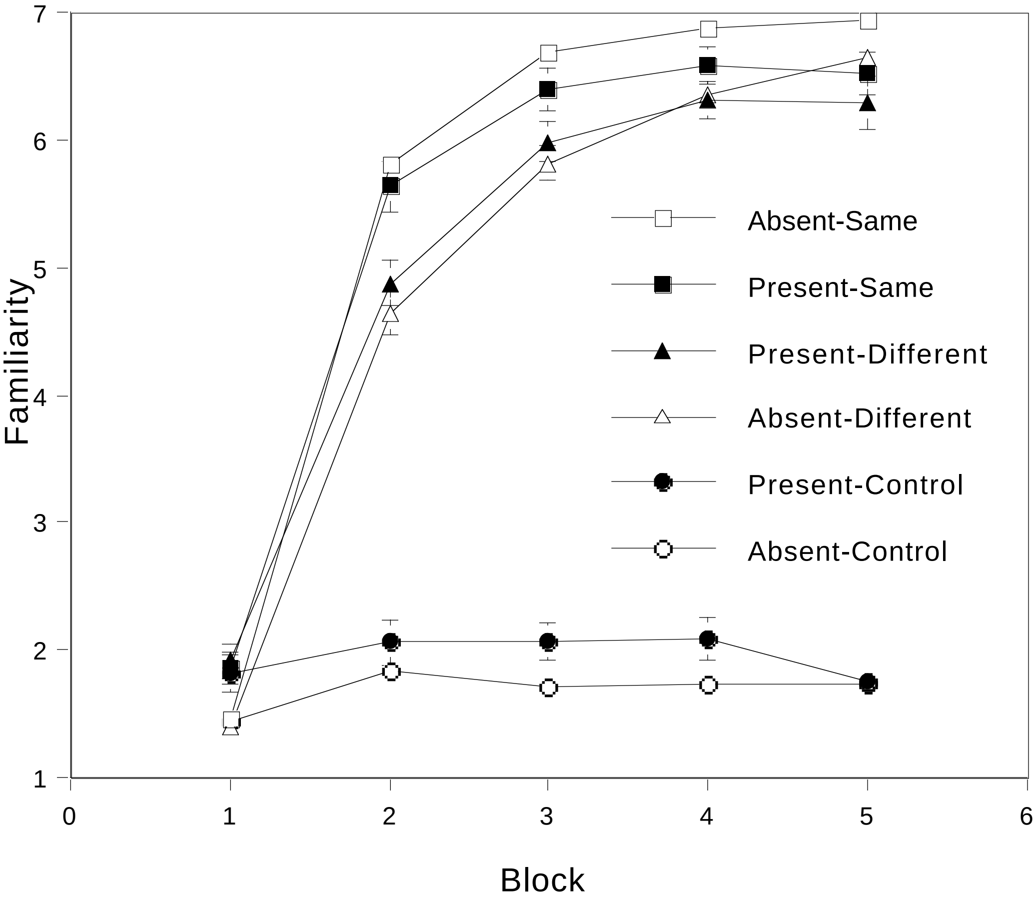


Figure 2.14. Comparison of the Effects of Viewpoint and Exposure on Familiarity Ratings in Experiments 1 and 2

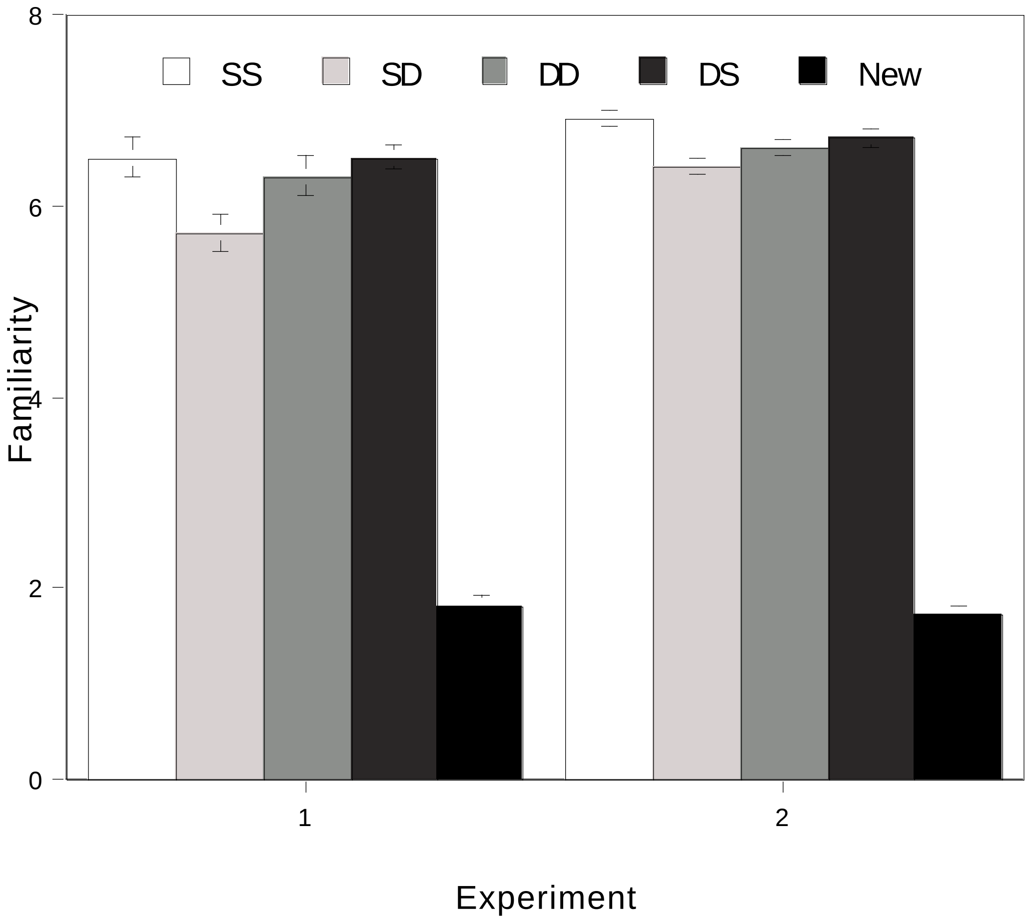


Figure 2.15. Familiarity Ratings in Block 5 of Experiments 1 and 2

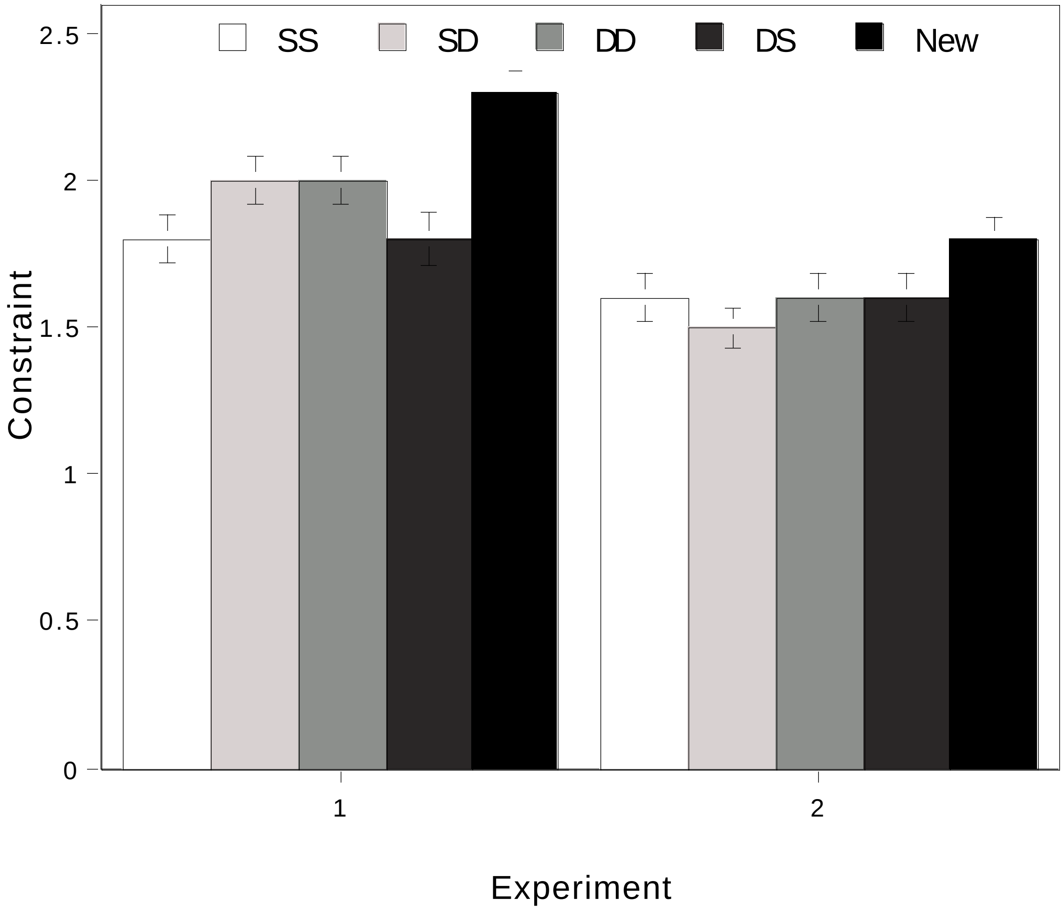


Figure 2.16. H1 Values in Experiments 1 and 2

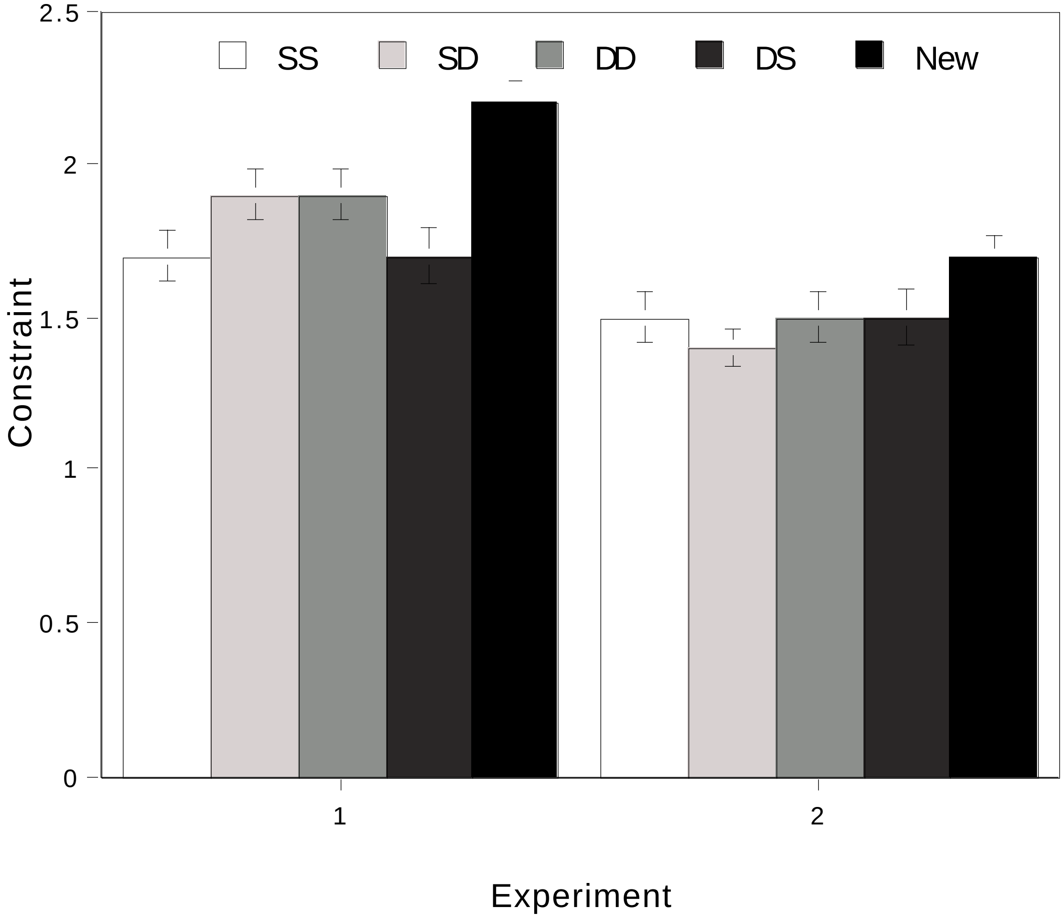


Figure 2.17. H1t Values in Experiments 1 and 2

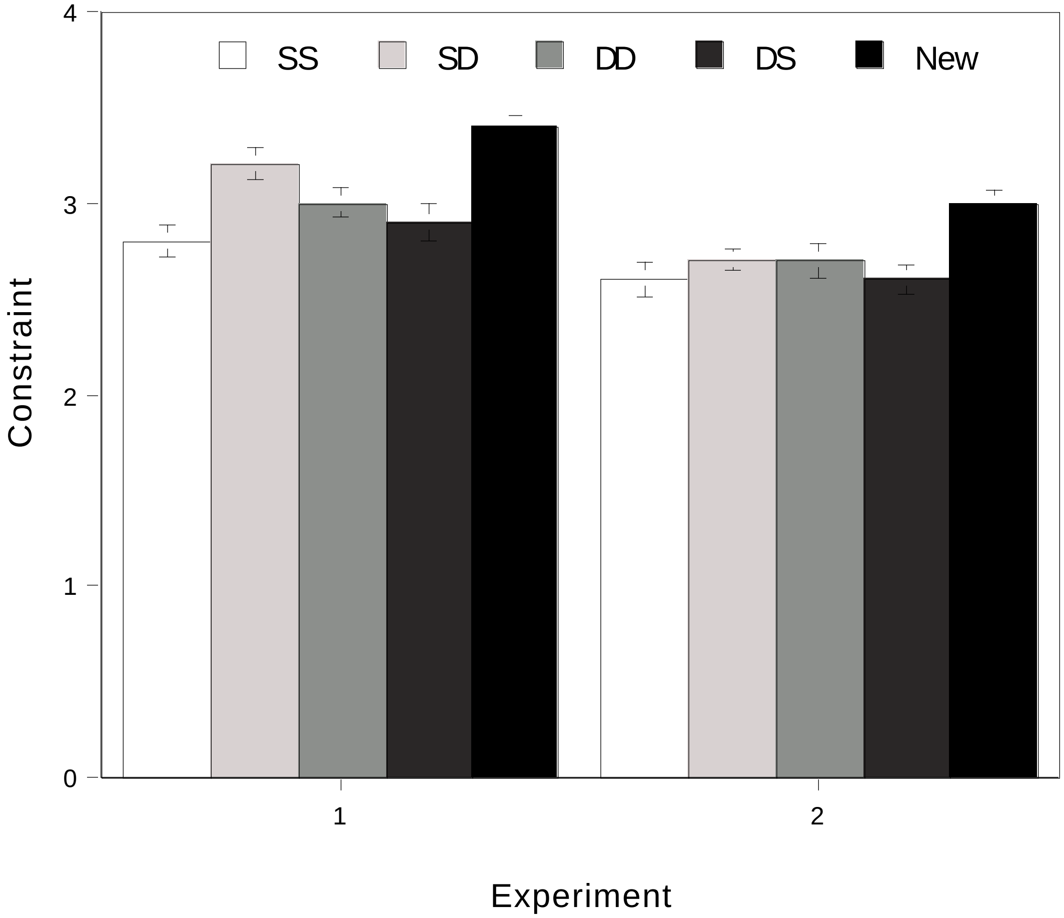


Figure 2.18. H2 Values in Experiments 1 and 2

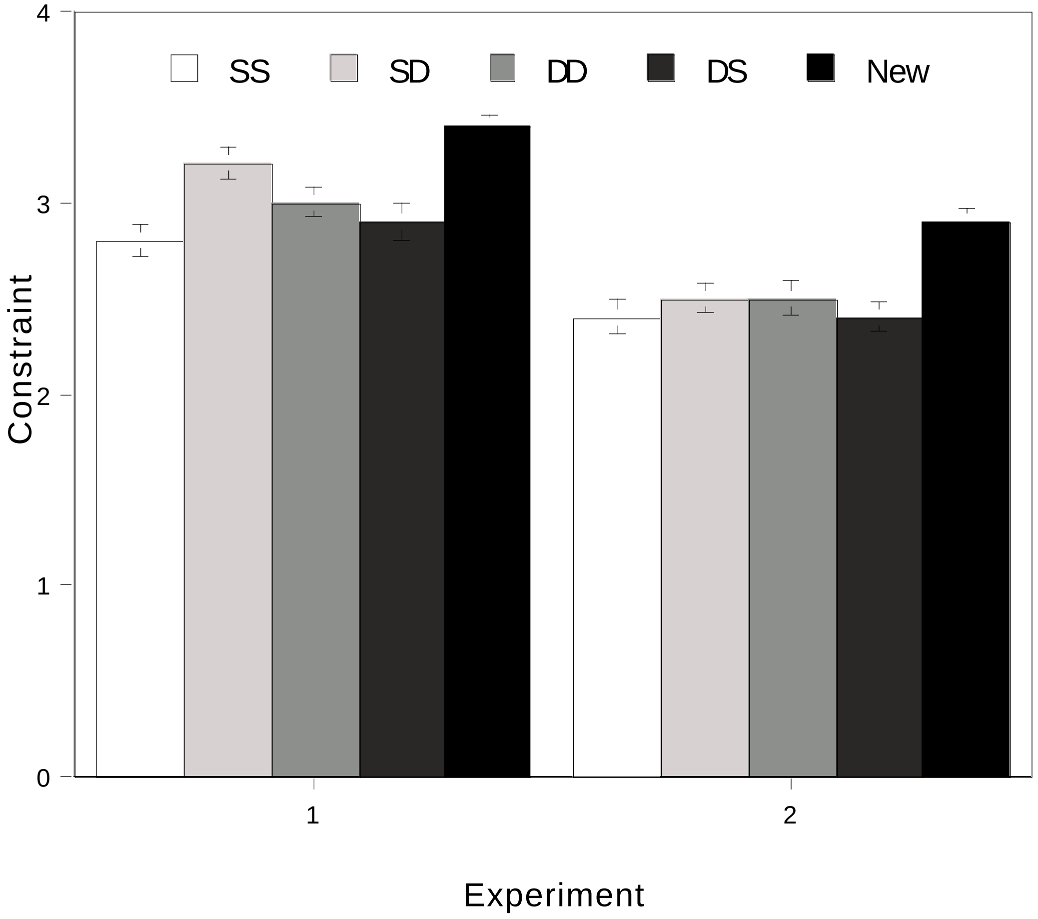


Figure 2.19. H2t Measure in Experiments 1 and 2

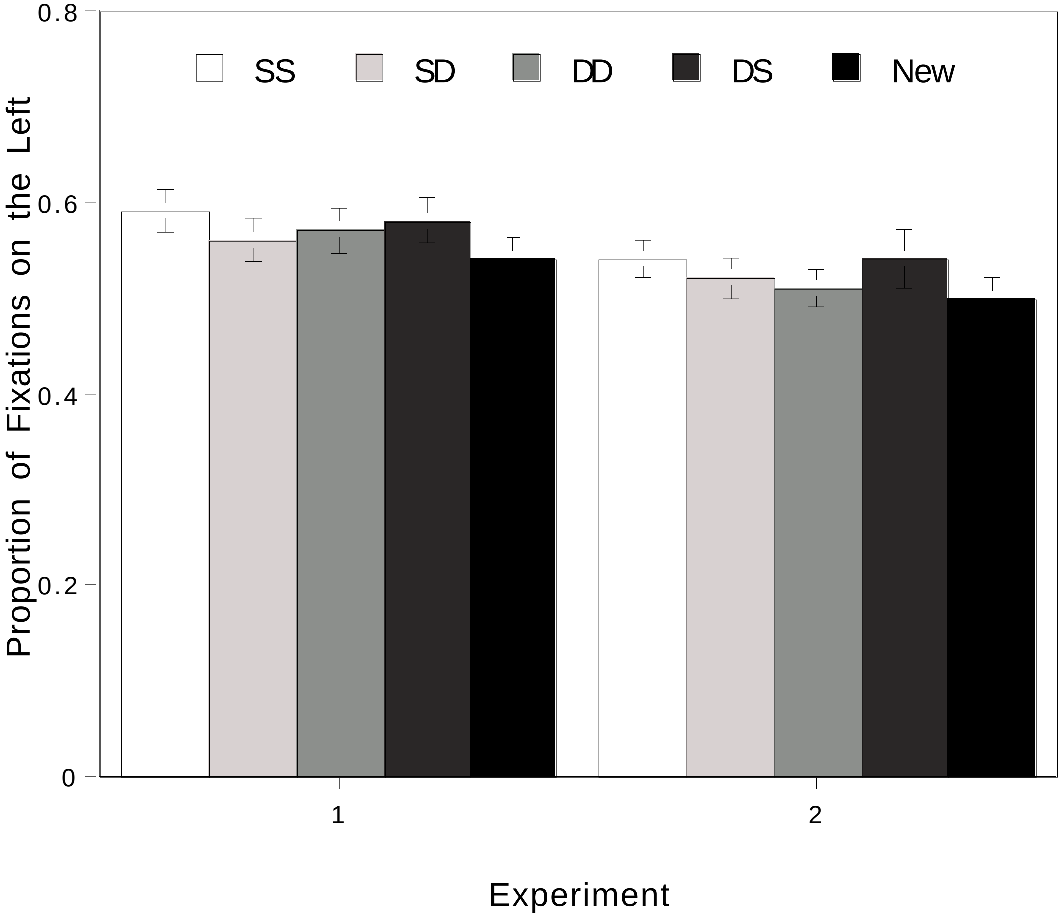


Figure 2.20. Proportion of Fixations on the Left Side of the Picture in Experiments 1 and 2

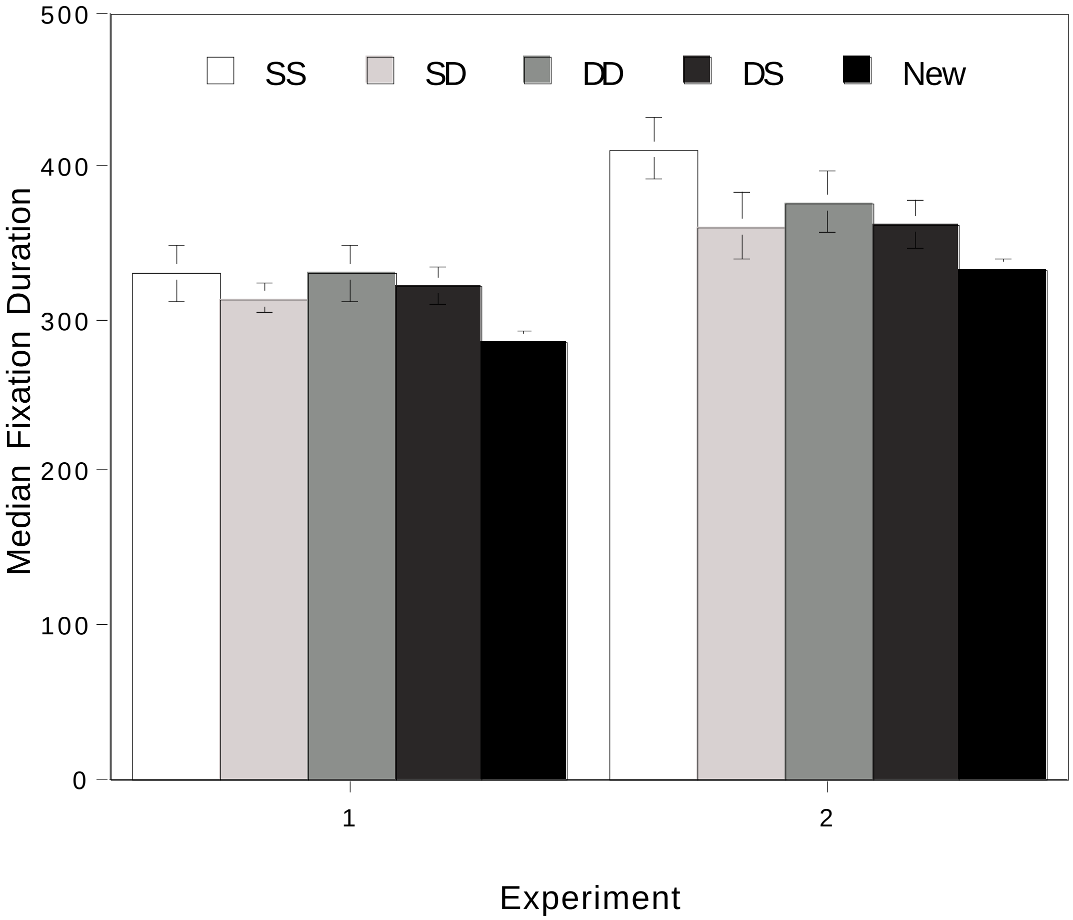


Figure 2.21. Median Fixation Durations in Experiments 1 and 2

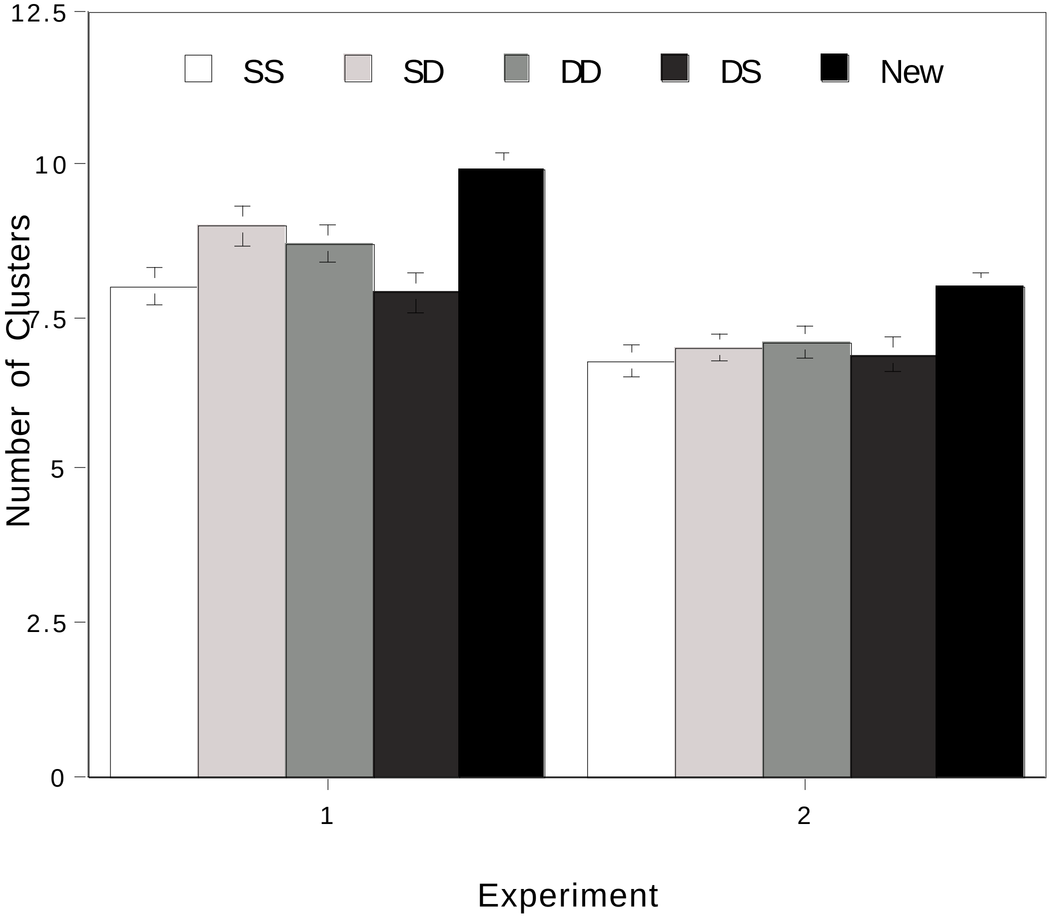


Figure 2.22. Number of Clusters in Experiments 1 and 2

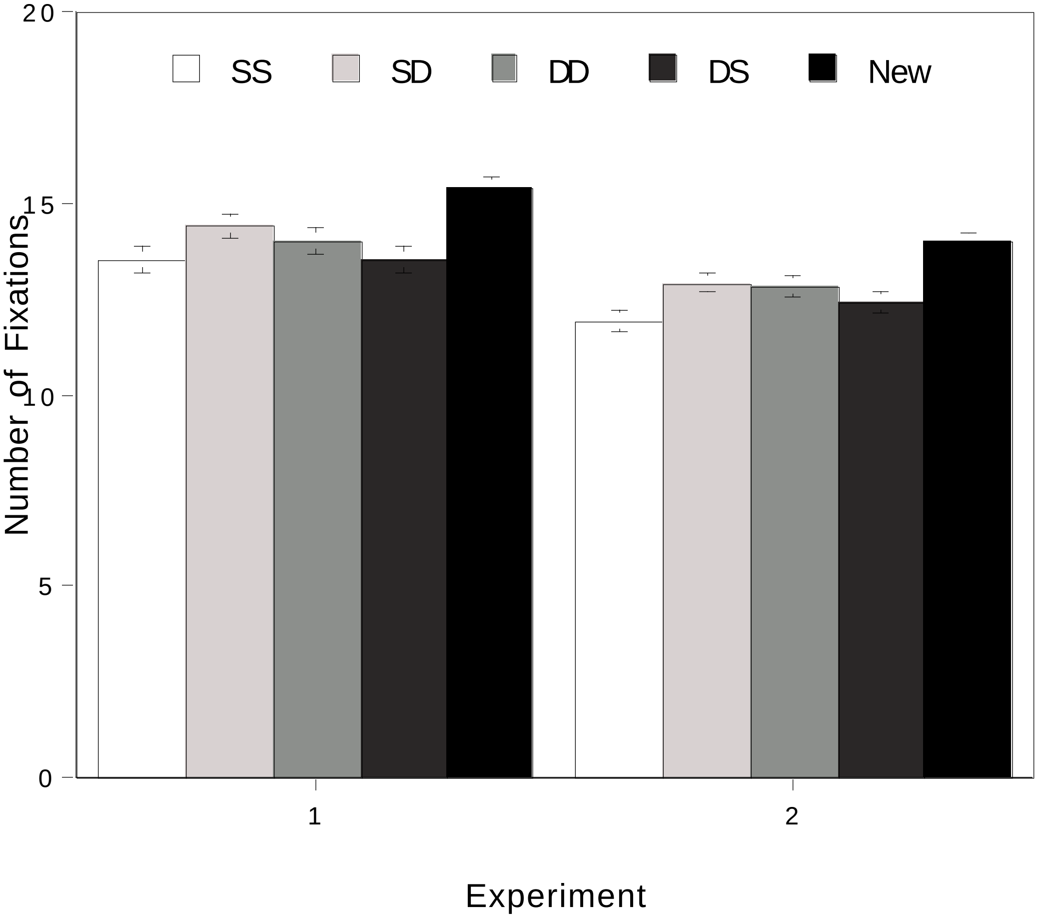


Figure 2.23. Number of Fixations in Experiments 1 and 2

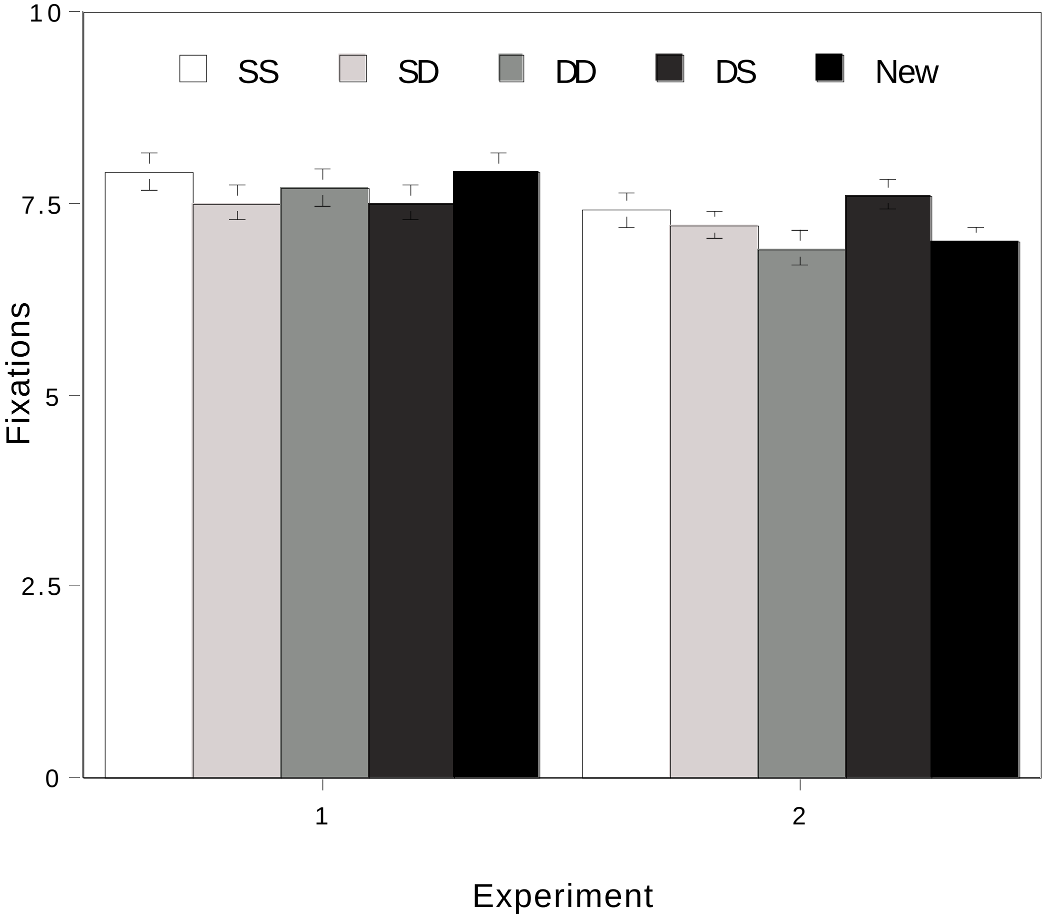


Figure 2.24. Number of Fixations to Return to Original Location in Experiments 1 and 2

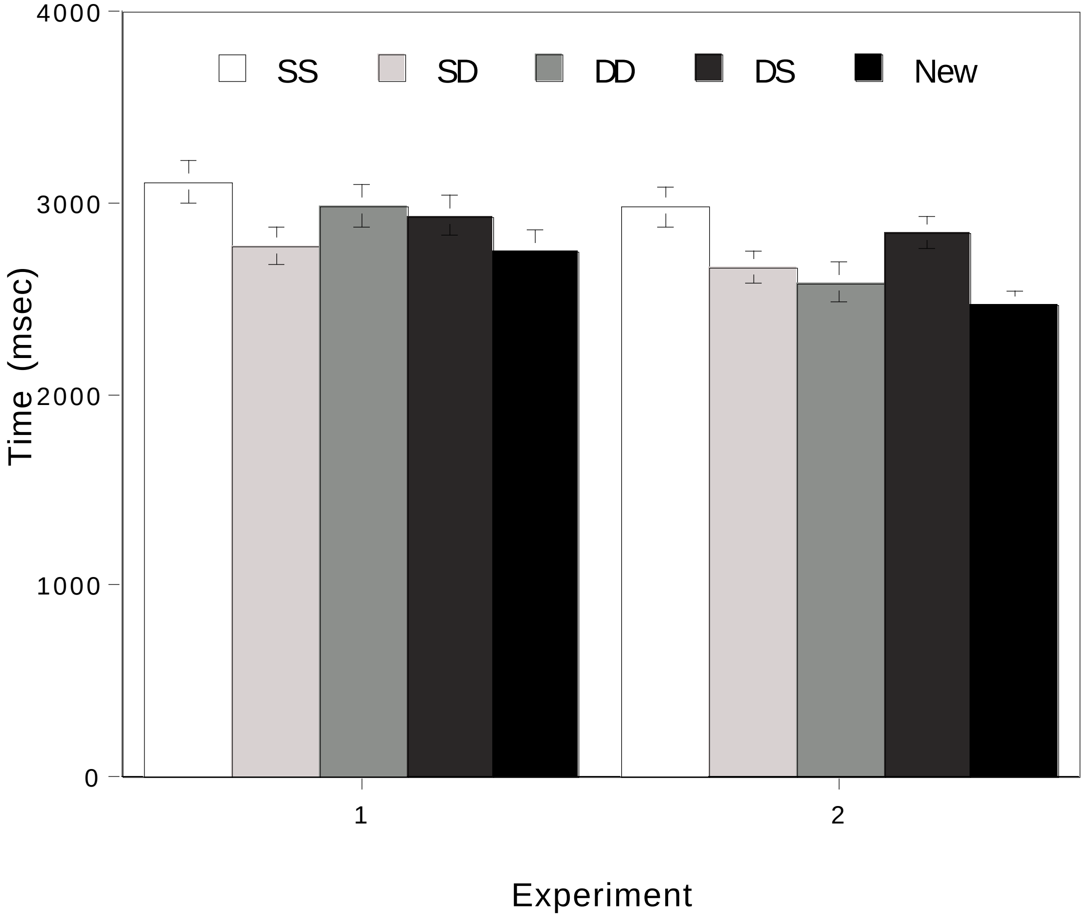


Figure 2.25. Time to Return to Original Location in Experiments 1 and 2

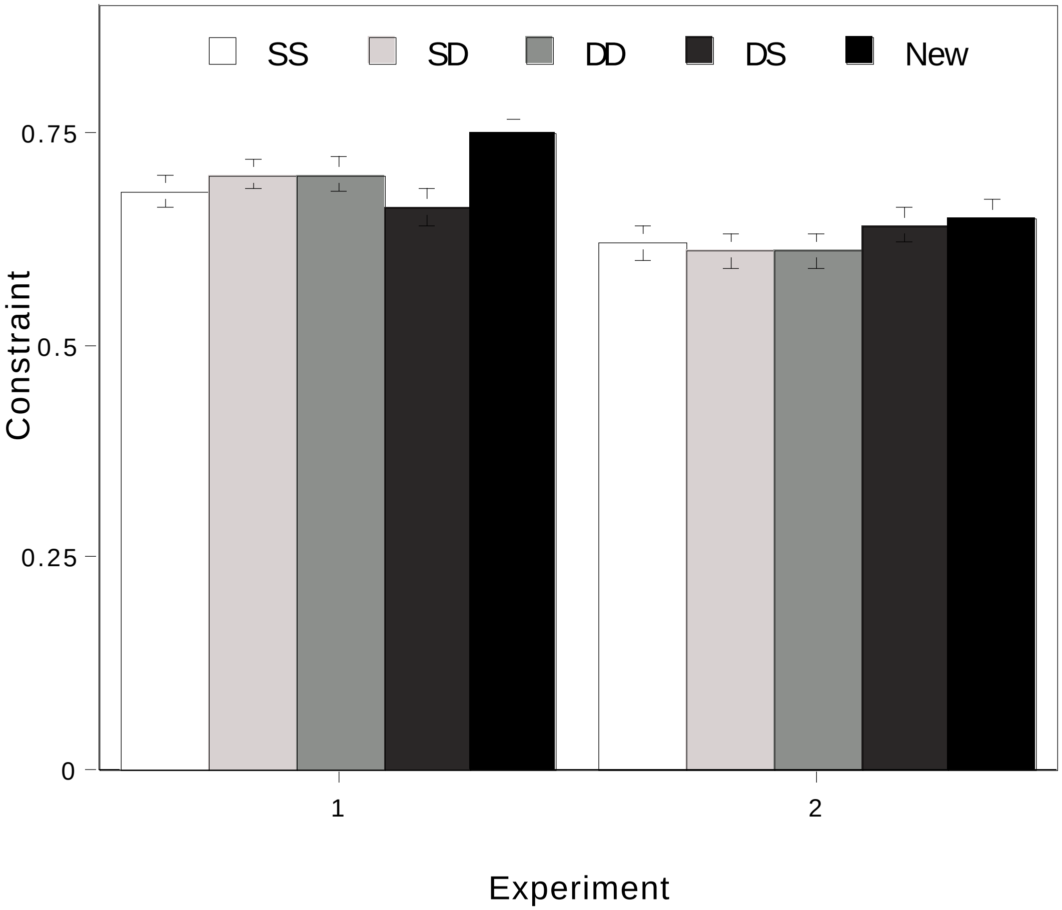


Figure 2.26. S1 Values in Experiments 1 and 2

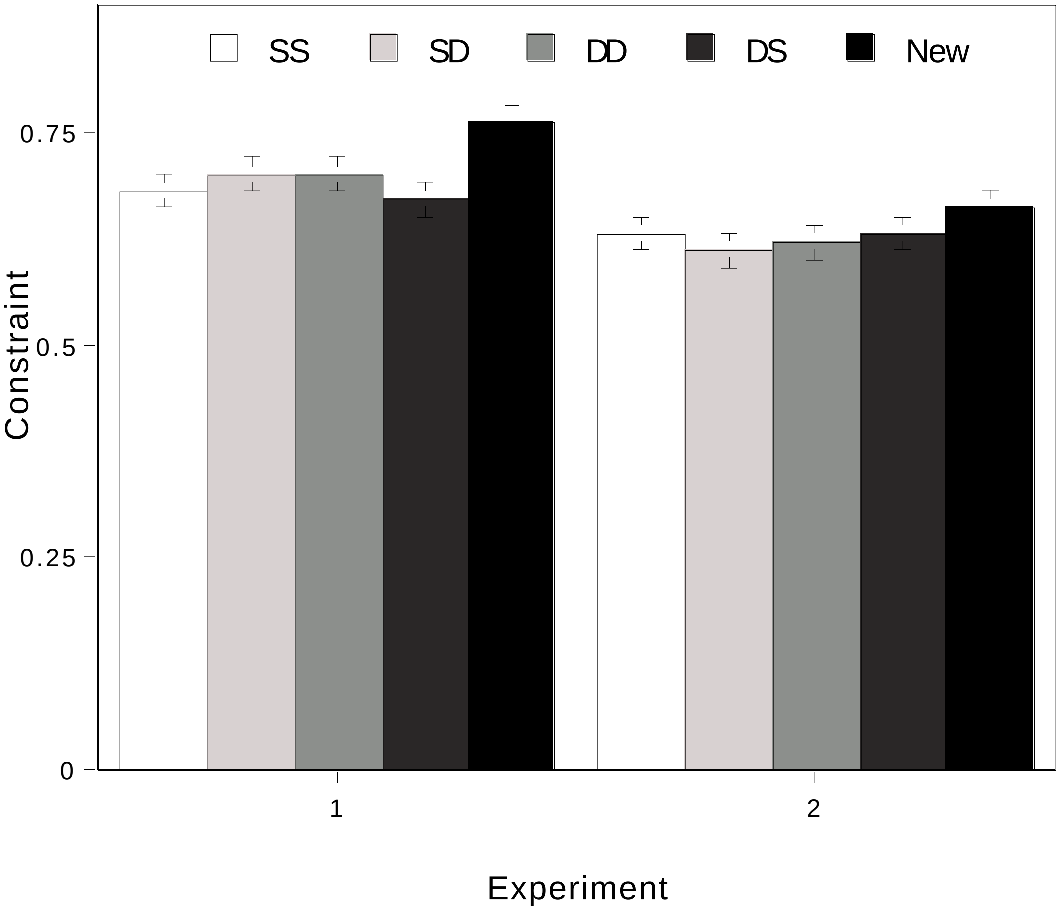


Figure 2.27. S1t Values in Experiments 1 and 2

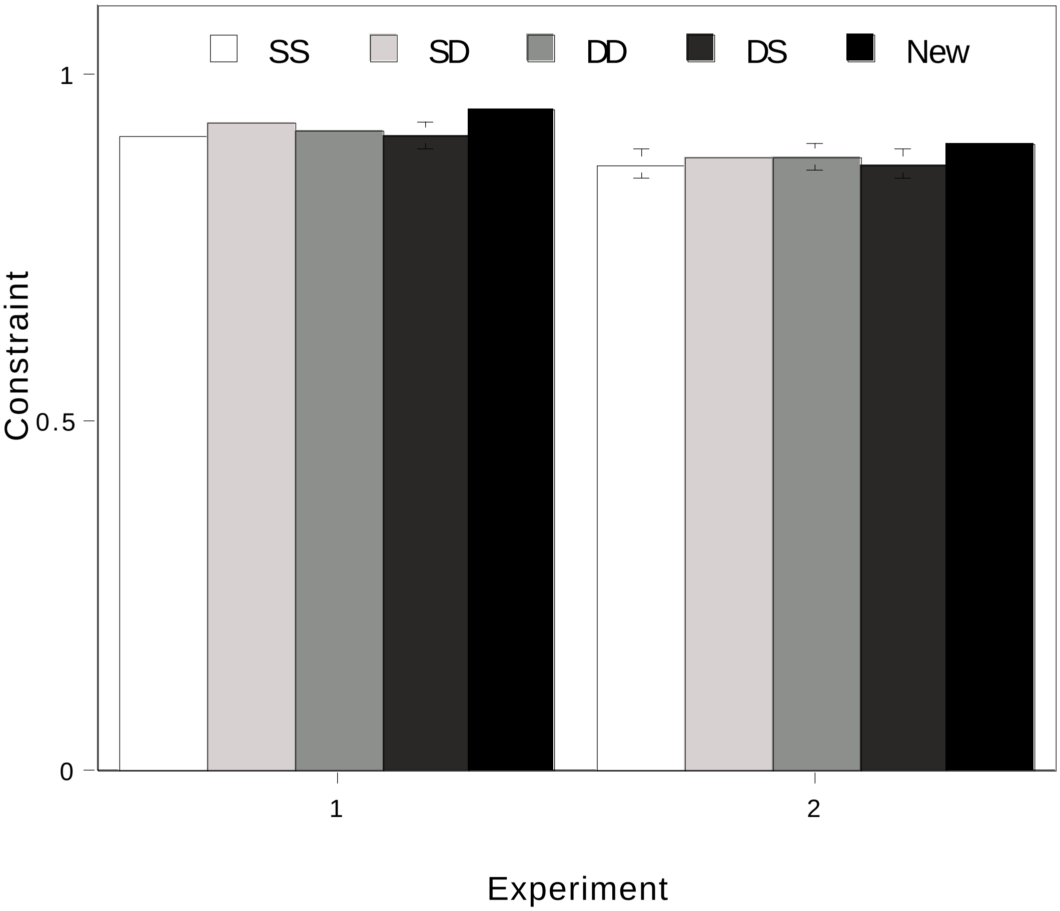


Figure 2.28. S2 Values in Experiments 1 and 2

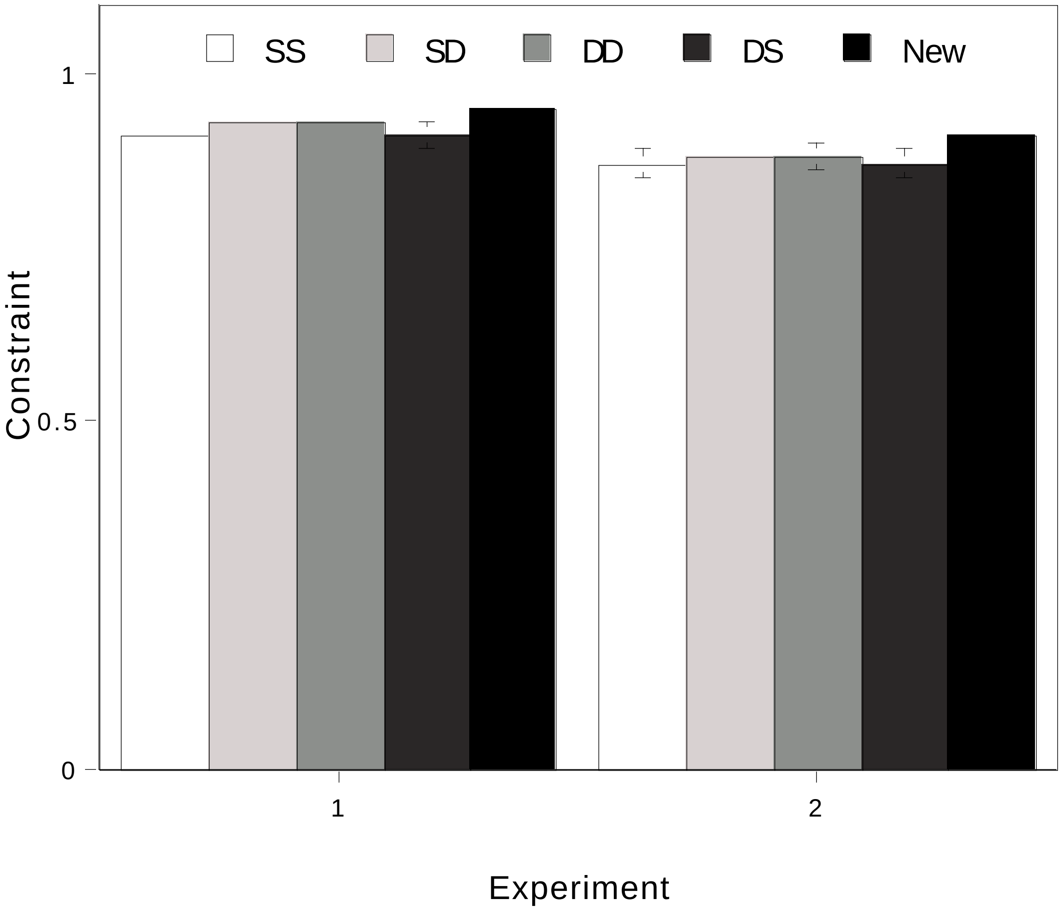


Figure 2.29. S2t Values in Experiments 1 and 2

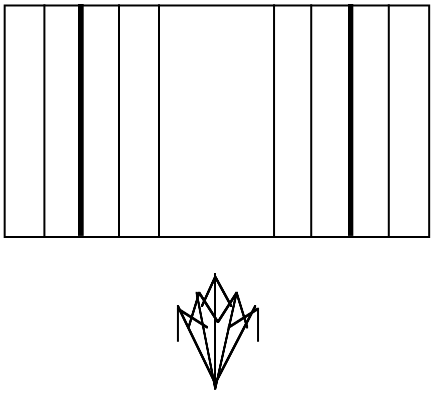


Figure 3.1. Schematic Depicting Viewer Perspective and Direction of Camera

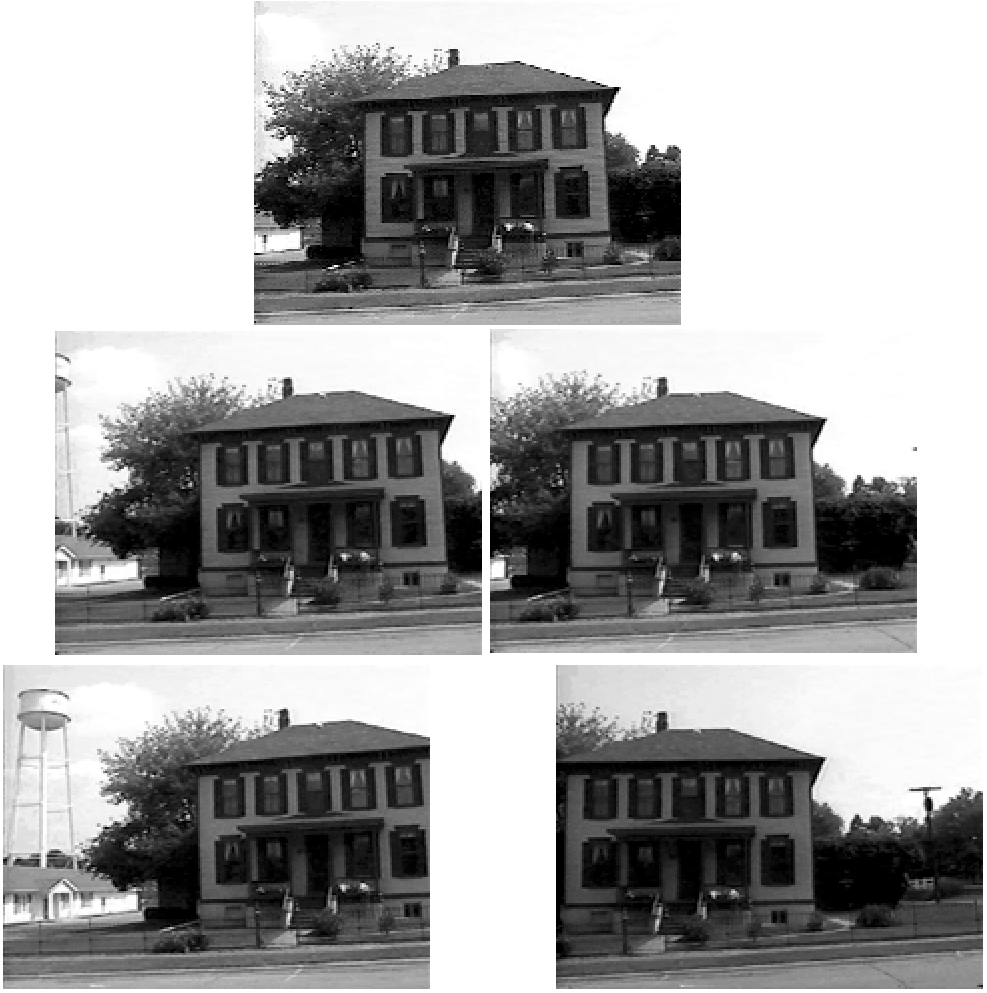


Figure 3.2. Examples of Stimuli Used in Experiment 3

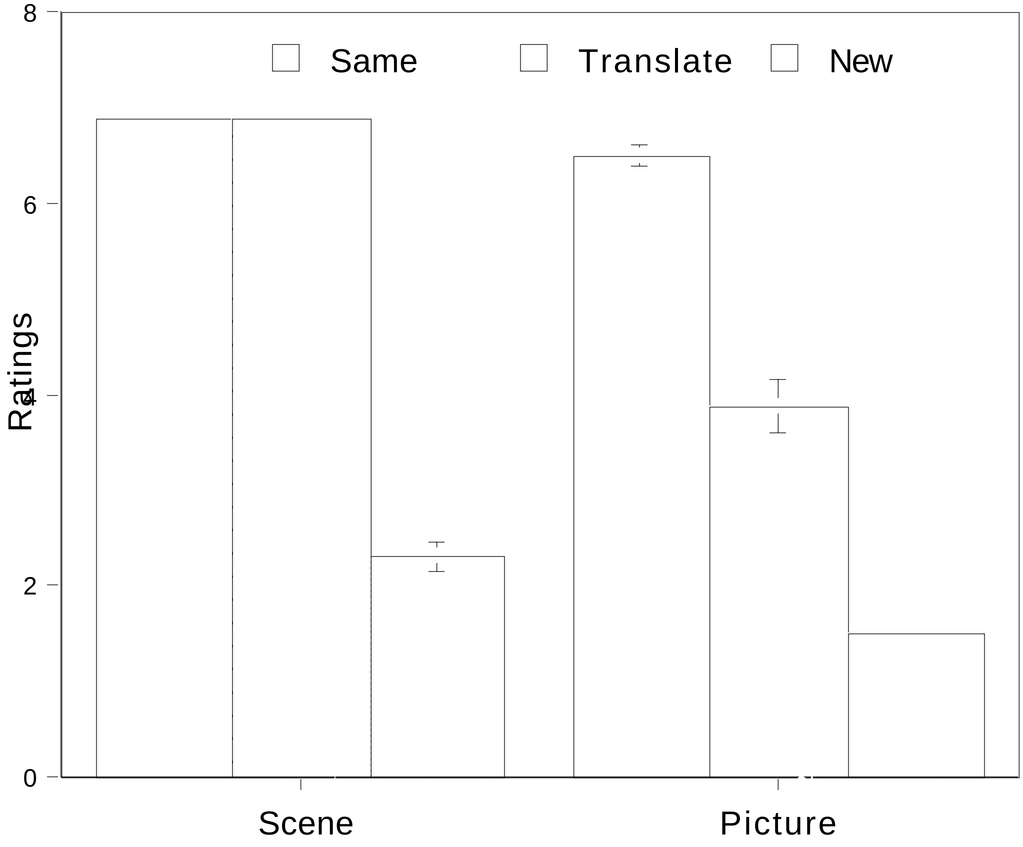


Figure 3.3. The Effect of Translation and Task on Recognition Confidence Ratings

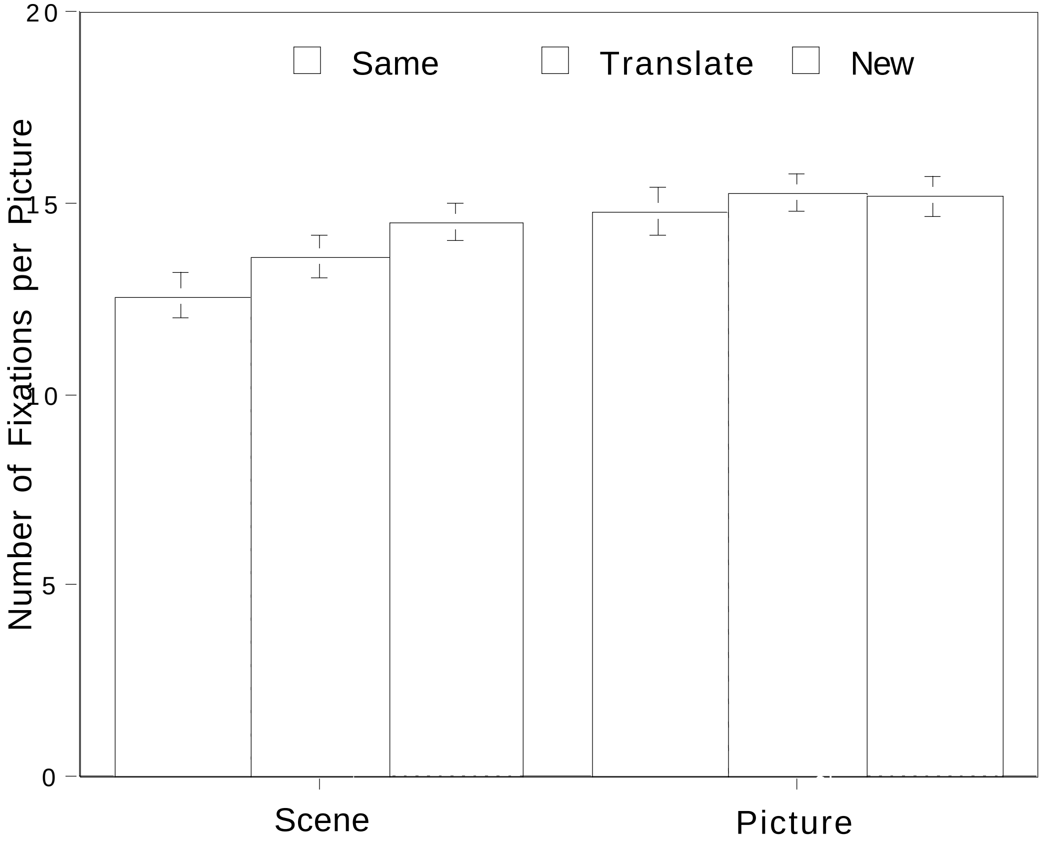


Figure 3.4. The Effect of Translation and Task on the Number of Fixations per Picture

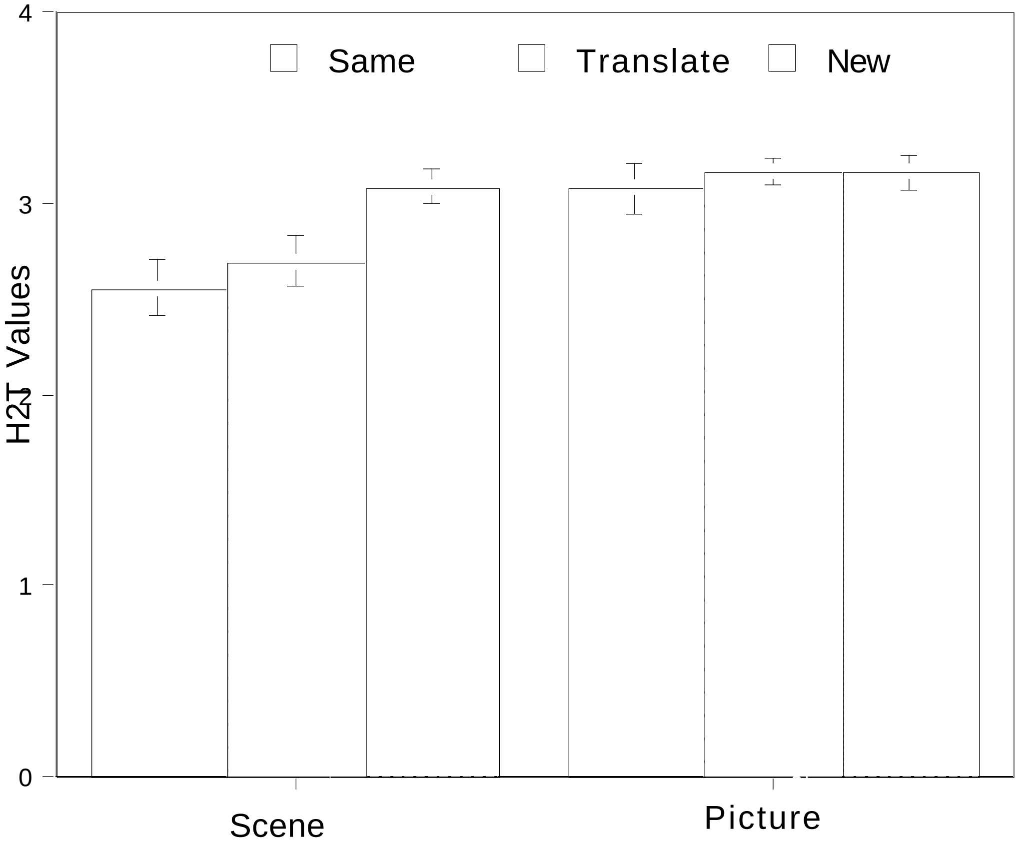


Figure 3.5. The Effect of Translation and Task on the H2T Entropy Measure

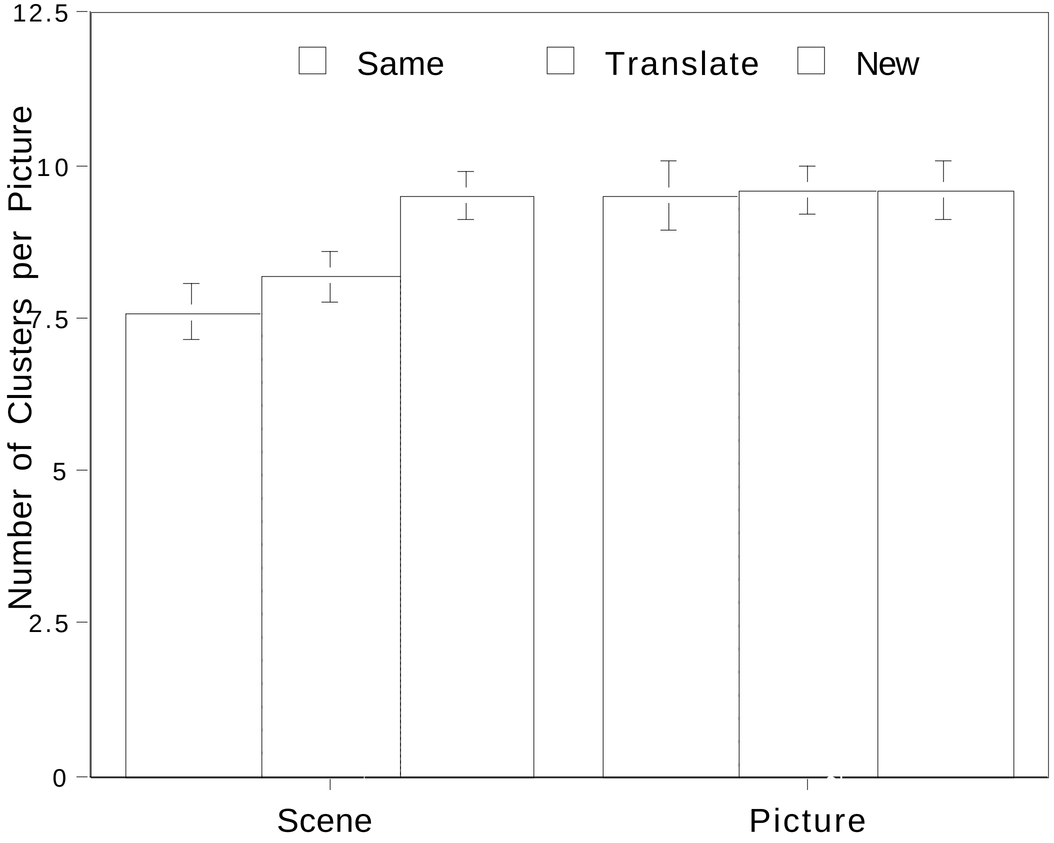


Figure 3.6. The Effect of Translation and Task on the Number of Clusters per Picture

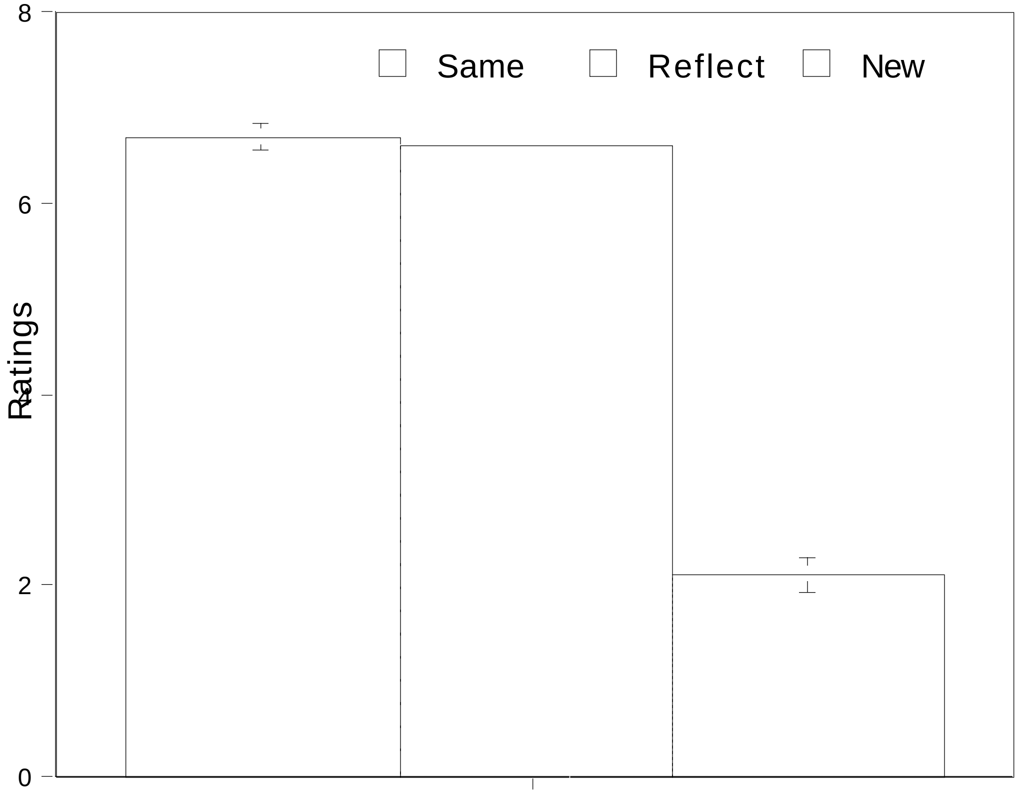


Figure 4.1. The Effect of Reflection on Recognition Confidence Ratings

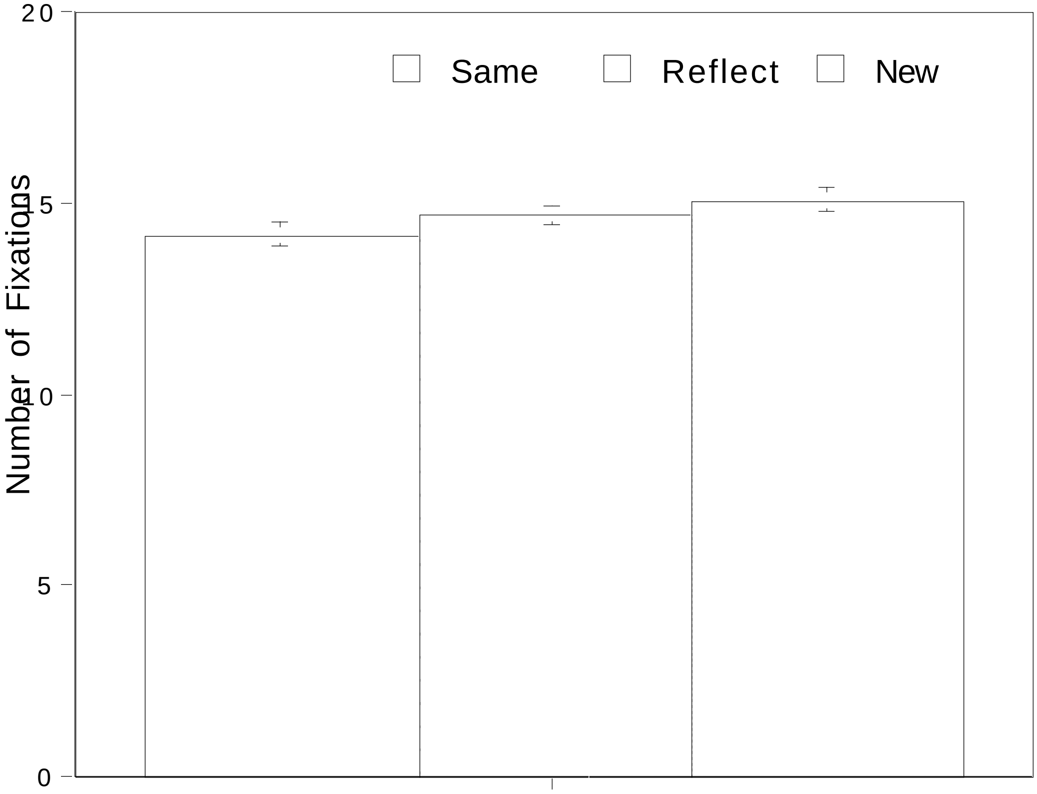


Figure 4.2. The Effect of Reflection on the Number of Fixations per Picture

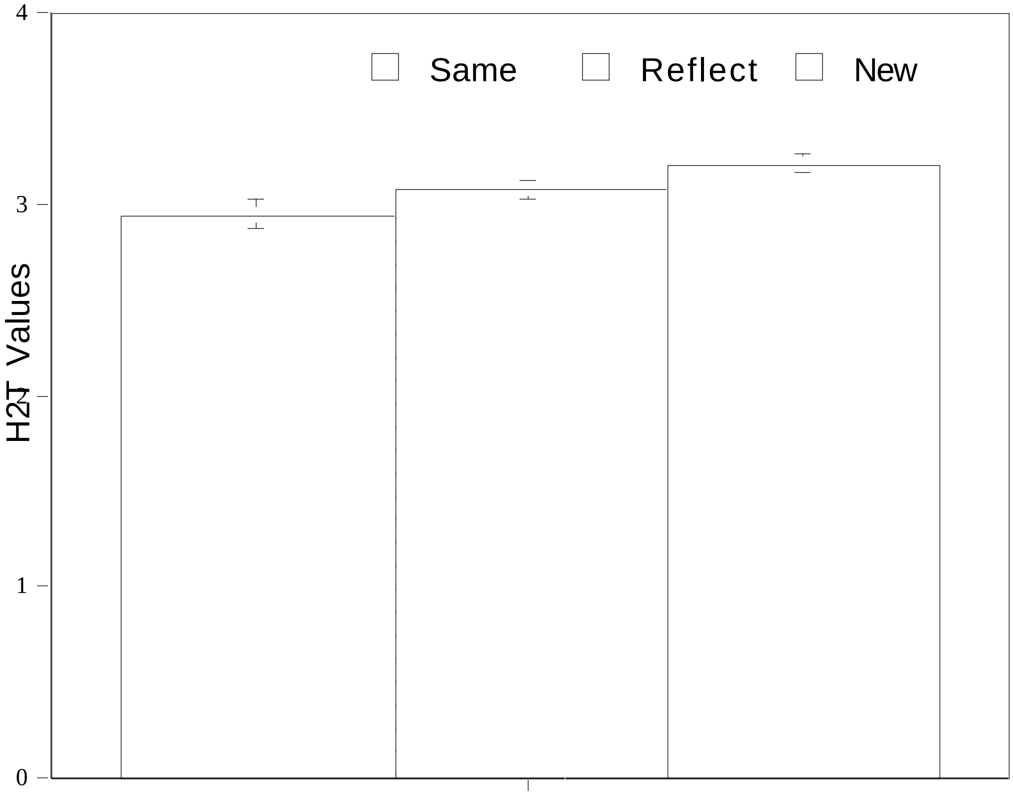


Figure 4.3. The Effect of Reflection on the H2T Entropy Measure

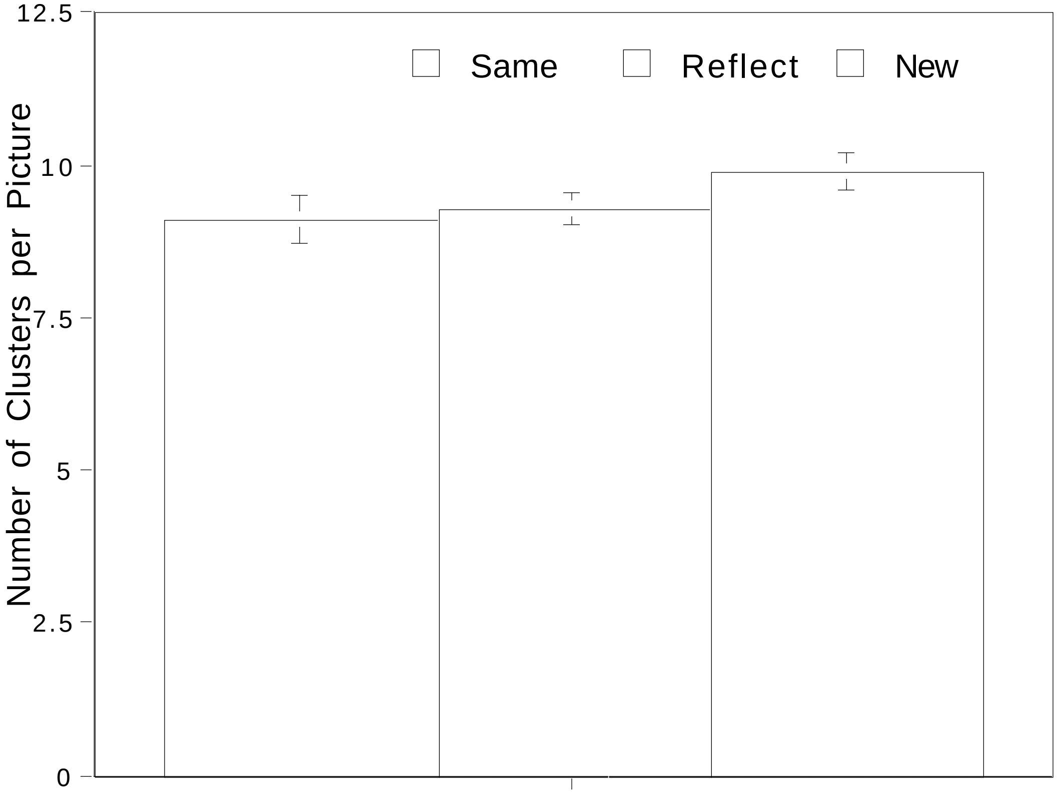


Figure 4.4. The Effect of Reflection on the Number of Clusters per Picture

Bibliography

Althoff, R.R. (1998). Eye movement-based memory assessment: the use of eye movement monitoring as an indirect measure of memory. Doctoral dissertation. University of Illinois at Urbana-Champaign.

Althoff, R.R., & Cohen, N.J. (in manuscript). Eye movement-based memory effect: A re-processing effect in face perception.

Antes, J. R. (1974). The time course of picture viewing. Journal of Experimental Psychology, 103(1), 62-70.

Bartlett, J.C., Gernsbacher, M.A., & Till, R.E. (1987). Remembering left-right orientation of pictures. Journal of Experimental Psychology: Learning, Memory, and Cognition, 13(1), 27-35.

Bartram, D.J. (1974). The role of visual and semantic codes in object naming. Cognitive Psychology, 6, 325-356.

Biederman, I., & Cooper, E. E. (1991). Evidence for complete translational and reflectional invariance in visual object priming. Perception, 20, 585-593.

Biederman, I., & Cooper, E.E. (1992). Size invariance in visual object priming. Journal of Experimental Psychology: Human Perception and Performance, 18, 121-133.

Biederman, I., & Gerhardstein, P. C. (1993). Recognizing depth-rotated objects: Evidence for 3D viewpoint invariance. Journal of Experimental Psychology: Human Perception and Performance, 19, 1162-1182.

Biederman, I., & Gerhardstein, P.C. (1995). Viewpoint-dependent mechanisms in visual object recognition: Reply to Tarr and Bulthoff (1995). Journal of Experimental Psychology:

Human Perception and Performance, 21, 1506-1504.

Brewer, W. F., & Treyens, J. C. (1981). Role of schemata in memory for places. Cognitive Psychology, 13, 207-230.

Bruyer, R. (1991). Indirect face recognition in prosopagnosia: a review. Brain and Cognition, 15, 223-235.

Bulthoff, H. H., & Edelman, S. (1992). Psychophysical support for a two-dimensional view-interpolation theory of object recognition. Proceedings of the National Academy of Science, 89, 60-64.

Bulthoff, H.H., Edelman, S., & Tarr, M.J. (1995). How are three-dimensional objects represented in the brain? Cerebral Cortex, 5, 247-260.

Carlson-Radvansky, L.A., & Irwin, D.E. (1994). Reference frame activation during spatial term assignment. Journal of Memory and Language, 33, 646-671.

Cave, C. B., & Squire, L. R. (1992). Intact and long-lasting repetition priming in amnesia. Journal of Experimental Psychology: Learning, Memory, and Cognition, 18(3), 509-520.

Cave, C.B., Bost, P.R., & Cobb, R.E. (1996). Effects of color and pattern on indirect and direct picture memory. Journal of Experimental Psychology: Learning, Memory, and Cognition, 22(3), 639-653.

Cohen, N.J.,& Eichenbaum, H. (1993). Memory, Amnesia, and the Hippocampal System. Cambridge, MA: MIT Press.

Cohen, N.J., Althoff, R.R, Webb, J.M., McConkie, G.W., Holden,J.A. & Noll, E.L. (submitted). Eye movement monitoring as an indirect measure of memory.

Cooper, L. A., Schacter, D. L., Ballesteros, S., & Moore, C. (1992). Priming and recognition of transformed three-dimensional objects: Effects of size and reflection. Journal of Experimental Psychology: Learning, Memory, and Cognition, 18(1), 43-57.

Diwadkar, V.A., & McNamara, T.P. (1997). Viewpoint dependence in scene recognition. Psychological Science, 8(4), 302-307.

Edelman, S., & Bulthoff, H. H. (1992). Orientation dependence in the recognition of familiar and novel views of three-dimensional objects. Vision Research, 12, 2385-2400.

Edelman, S., & Weinshall, D. (1991). A self-organizing multiple-view representation of 3D objects. Biological Cybernetics, 64, 209-219.

Farah, M.J. (1991). Patterns of co-occurence among the associative agnosias: Implications for visual object representation. Cognitive Neuropsychology, 8(1), 1-19.

Farah, M. J., Levinson, K. L., & Klein, K. L. (1995). Face perception and within-category discrimination in prosopagnosia. Neuropsychologica, 33(6), 661-674.

Farah, M.J., Tanaka, J.W., & Drain, H.M. (1995). What causes the face inversion effect? Journal of Experimental Psychology: Human Perception and Performance, 21(3), 628-634.

Feinberg, T. E., Dyckes-Berke, D., Miner, C., & Roane, D. M. (1995). Knowledge, indirect knowldge and metaknowledge in visual agnosia and pure alexia. Brain, 118, 789-800.

Fiser, J., & Biederman, I. (1995). Size invariance in visual object priming of grey-scale images. Perception, 24, 741-748.

Frederickson, R.E. & Bartlett, J.C. (1987). Cognitive impenetrability of memory for orientation. Journal of Experimental Psychology: Learning, Memory, and Cognition, 13(2), 269-277.

Gautheir, I., & Tarr, M.J. (1997). Orientation priming of novel shapes in the context of viewpoint-dependent recognition. Perception, 26, 51-73.

Gernsbacher, M.A. (1985). Surface information loss in comprehension. Cognitive Psychology, 17, 324-363.

Graf, P.,& Schacter, D.I. (1985). Implicit and explicit memory for new associations in normal and amnesic subjects. Journal of Experimental Psychology: Learning, Memory, and Cognition, 11, 501-518.

Hayward, W.G. & Tarr, M.J. 91997). Testing conditions for viewpoint invariance in object recognition. Journal of Experimental Psychology: Human Perception and Performance, 23(5), 1511-1521.

Hill, H., Bruce, V., & Akamatsu, S. (1995). Perceiving the sex and race of faces: the role of shape and color. Proceedings of the Royal Society of London, 261, 367-373.

Hummel, J. E., & Biederman, I. (1992). Dynamic binding in a neural network for shape recognition. Psychological Review, 99(3), 480-517.

Humphreys, G.K., & Kahn, S.C. (1992). Recognizing novel views of three-dimensional objects. Canadian Journal of Psychology, 46, 170-190.

Humphreys, G. K., Goodale, M. A., Jakobson, L. S., & Servos, P. (1994). The role of surface information in object recognition: studies of a visual form agnosic and normal subjects. Perception, 23, 1457-1481.

Humphreys, G. W., & Riddoch, M. J. (1984). Routes to object constancy: Implications from neurological impairments of object constancy. Quarterly Journal of Experimental Psychology, 36(A), 385-415.

Intraub, H., & Richardson, M. (1989). Wide-Angle memories of close-up scenes. Journal of Experimental Psychology: Learning, Memory, and Cognition, 15(2), 179-187.

Intraub, H., & Berkowits, D. (1996). Beyond the edges of a picture. American Journal of Psychology, 109(4), 581-598.

Intraub, H., Gottesman, C.V., Willey, E.V., & Zuk, I.J. (1996). Boundary extension for briefly glimpsed photographs: Do common perceptual processes result in unexpected memory distortions? Journal of Memory and Language, 35, 118-134.

Jolicoeur, P. (1987). A size-congruency effect in memory for visual shape. Memory and Cognition, 15, 531-543.

Jones, G.V. (1990). Misremembering a familiar object: when left is not right. Memory and Cognition, 18, 174-182.

Kendrick , K. M., Atkins, K., Hinton, M. R., Broad, K. D., Fabre-Nys, C., & Keverne, B. (1995). Facial and vocal discrimination in sheep. Animal Behavior, 49, 1665-1676.

Klatsky, R.L., & Forrest, F.H. (1984). Recognizing familiar and unfamilar faces. Memory and Cognition, 12, 60-70.

Kolers, P.A., Dunchinicky, R.L., & Sundstroem, G. (1985). Size in the visual processing of faces and words. Journal of Experimental Psychology: Human Perception and Performance, 11, 726-751.

Kosslyn, S. M., Alpert, N. M., Thompson, W. L., Chabris, C. F., Rauch, S. L., & Anderson, A. K. (1994). Identifying objects seen from different viewpoints: a PET investigation. Brain, 117, 1055-1071.

Lawson, R., & Humphreys, G.W. (1996). View specificity in object processing: Evidence from picture matching. Journal of Experimental Psychology: Human Perception and

Performance, 27(2), 395-416.

Levelt, W.J.M. (1989). Speaking: From Intention to Articulation. Cambridge: MIT Press.

Logothetis, N. K., & Pauls, J. (1995). Psychophysical and physiological evidence for viewer-centered object representations in the primate. Cerebral Cortex, 3, 270-288.

Logothetis, N. K., Pauls, J., Bulthoff, H. H., & Poggio, T. (1994). View-dependent object recognition by monkeys. Current Biology, 4(5), 401-414.

Logothetis, N. K., Pauls, J., & Poggio, T. (1995). Shape representation in the inferior temporal cortex of monkeys. Current Biology, 5(5), 552-563.

Logothetis, N.K., & Sheinberg, D.L. (1996). Visual object recognition. Annual Review of Neuroscience, 19, 577-621.

Lynch, J.C, Graybiel, A.M., & Lobeck, L.J. (1985). The differential projection of two cytoarchitectural subregions of the inferior parietal lobule of macaque upon the deep layers of the superior colliculus. Journal of Computational Neurology, 235, 242-254.

Maki, R.H. & Marek, M.N. (1997). Egocentric spatial framework effects from single and multiple points of view. Memory and Cognition, 25(5), 677-690.

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

Martin, M., & Jones, G.V. (1995). Danegeld remembered: taxing further the coin head illusion. Memory, 3, 97-104.

Martin, M., & Jones, G.V. (1997). Memory for orientation in the natural environment. Applied Cognitive Psychology, 11, 279-288.

McKelvie, S. (1983). Effects of lateral reversal on recognition memory for photographs of faces. British Journal of Psychology, 74, 391-407.

McNamara, T.P. (1986). Mental representations of spatial relations. <u>Cognitive Psychology, 18</u>, 87-121.

Maciukenas, M.A., Althoff, R.R., Holden, J.A., Webb, J.M., & Cohen, N.J. (1997). EMTool (Version 2.0) [Computer software]. Urbana, IL: Amnesia Research Laboratory.

Milliken, B., & Jolicoeur, P. (1992). Size effects in visual recognition memory are determined by perceived size. <u>Memory and Cognition, 20</u>, 83-95.

Missal, M., Vogels, R., & Orban, G.A. (1997). Responses of macaque inferior temporal neurons to overlapping shapes. <u>Cerebral Cortex, 7</u>, 758-767.

Miyashita, Y. (1993). Inferior temporal cortex: Where visual perception meets memory. <u>Annual Review of Neuroscience, 16</u>, 245-263.

Molnar, F. & Ratsikas, D. (1987). Some aesthetical aspects of visual exploration. In O'Reagan, J.K. & Levy-Schoen, A. (Eds.), <u>Eye Movement: From Physiology to Cognition</u>. North Holland: Elsevier.

Perrett, D. I., Smith, P. A. J., Potter, D. D., Mistlin, A. J., Head, A. S., Milner, A. D., & Jeeves, M. A. (1985). Visual cells in the temporal cortex sensitive to face view and gaze direction. <u>Proceedings of the Royal Society of London, 223</u>, 293-317.

Poggio, T., & Edelman, S. (1990). A network that learns to recognize three-dimensional objects. <u>Nature, 343</u>, 263-266.

Price, C. J., & Humphreys, G. W. (1989). The effects of surface detail on object categorization and naming. <u>The Quarterly Journal of Experimental Psychology, 41</u>(4), 797-828.

Puce, A., Allison, T., Gore, J.C., McCarthy, G. (1995). Face-sensitive regions in human extrastriate cortex studied by functional MRI. <u>Journal of Neurophysiology, 74</u>(3), 1192-1199.

Quinlan, P. T. (1991). Differing approaches to two-dimensional shape recognition. Psychological Bulletin, 109(2), 224-241.

Rapcsak, S.Z., Polster, M.R., Gilsky, M.L., Comer, J.F. (1996). False recognition of unfamiliar faces following right hemisphere damage: Neuropsychological and anatomical observations. Cortex, 32, 593-611.

Richardson-Klavehn. A., & Bjork, R.A. (1988). Measures of memory. Annual Review of Psychology, 39, 475-543.

Rizzo, M., Hurtig, R., & Damasio, A. R. (1987). The role of scanpaths in facial recognition and learning. Annals of Neurology, 22, 41-45.

Rolls, E. T., Tovee, M. J., Purcell, D. G., Stewart, A. L., & Azzopardi, P. (1994). The responses of neurons in the temporal cortex of primates in face identification and detection. Brain Research, 101, 473-484.

Sanocki, T., & Epstein, W. (1997). Priming the spatial layout of scenes. Psychological Science, 8(5), 374-378.

Schacter, D.I. (1987). Implicit memory: History and current status. Journal of Experimental Psychology: Learning, Memory, and Cognition, 13, 501-518

Schacter, D.I., & Cooper, L.A. (1993). Indirect and direct memory for novel visual objects: Structure and function. Journal of Experimental Psychology: Learning, Memory, and Cognition, 19, 995-1009.

Schacter, D.L., Cooper, L.A., & Delaney, S.M. (1990). Indirect memory for unfamiliar objects depends on access to structural descriptions. Journal of Experimental Psychology: General, 119, 5-24.

Schacter, D.L., Cooper, L.A., Delaney, S.M., Petersen, M.A., & Tharan, M. (1991). Indirect memory for possible and impossible objects: Constraints on the construction of structural descriptions. Journal of Experimental Psychology: Learning, Memory, and Cognition, 17, 3-19.

Schall, J. D., Morel, A., King, D. J., & Fullier, J. (1995). Topography of visual cortex connections with frontal eye fields in macaque: convergence and segregation of processing streams. Journal of Neuroscience, 15(6), 4464-4487.

Schiller, P. H. (1995). Effect of lesions in visual cortical area V4 on the recognition of transformed objects. Nature, 376, 342-344.

Seamon, J.G., Ganor-Stern, D., Crowley, M.J., Wilson, S.M., Weber, W.J., O'Rourke, C.M., & Mahoney, J.K. (1997). A mere exposure effect for transformed three-dimensional objects: Effects of reflection, size, or color changes on affect and recognition. Memory and Cognition, 25(3), 367-374.

Sergent, J., Ohta, S., & Macdonald, B. (1992). Functional neuroanatomy of face and object processing. Brain, 115, 15-36.

Sergent, J., & Signoret, J.L. (1992). Varieties of functional deficits in prosopagnosia. Cerebral Cortex, 2, 375-388.

Shapiro, P. N., & Penrod, S. (1986). Meta-analysis of facial identification studies. Psychological Bulletin, 100(2), 139-156.

Shepard, R.N., & Hurwitz, S. (1984). Upward direction, mental rotation, and discrimination of left and right turns in maps. Cognition, 18, 161-193.

Sholl, M.J., & Nolin, T.L. (1997). Orientation specificity in representation of place.

Journal of Experimental Psychology: Learning, Memory, and Cognition, 23(6), 1494-1507.

Srinivas, K. (1993). Perceptual Specificity in nonverbal priming. Journal of Experimental Psychology: Learning, Memory, and Cognition, 19(3), 582-602.

Srinivas, K. (1995). Representation of rotated objects in direct and indirect memory. Journal of Experimental Psychology: Learning, Memory, and Cognition, 21(4), 1019-1036.

Srinivas, K. (1996a). Size and reflection effects in priming: A test of transfer-appropriate processing. Memory and Cognition, 24(4), 441-452.

Srinivas, K. (1996b). Contrast and illumination effects on direct and indirect measures of memory. Journal of Experimental Psychology: Learning, Memory, and Cognition, 22(5), 1123-1135.

Standing, L., Conezio, J., & Haber, R.N. (1970). Perception and memory for pictures: single-trial learning of 2500 visual stimuli. Psychonomic Society, 19(2), 73-74.

Tanaka, K. (1993). Neuronal mechanisms of object recognition. Science, 262, 685-688.

Tarr, M.J. (1989). Orientation dependence in three-dimensional object recognition. Unpublished doctoral dissertation. MIT, Cambridge, MA.

Tarr, M.J. (1995). Rotating objects to recognize them: A case study of the role of viewpoint dependency in the recognition of three-dimensional objects. Psychonomic Bulletin and Review, 2, 55-82.

Tarr, M.J., & Bulthoff, H.H. (1995). Is human object recognition better described by geon-structural-descriptions or by multiple views? Journal of Experimental Psychology: Human Perception and Performance, 21, 1494-1505.

Tarr, M.J., Bulthoff, H.H., Zabinski, M., & Blanz, V. (1997). To what extent do unique parts influence recognition across changes in viewpoint? Psychological Science, 8(4), 282-289.

Taylor, H.A., & Tversky, B. (1992). Descriptions and depictions of environments. Memory and Cognition, 20(5), 483-496.

Taylor, H.A., & Tversky, B. (1996). Perspectives in spatial descriptions. Journal of Memory and Language, 35(3), 371-391.

Thorndyke, P.W., & Hayes-Roth, B. (1982). Differences in spatial knowledge acquired from maps and navigation. Cognitive Psychology, 14(4), 560-589.

Ullman, S. (1989). Aligning pictorial descriptions: an approach to object recognition. Cognition, 32, 193-254.

Van Lier, R.J., Leeuwenberg, E.L.J., & van der Helm, P.A. (1997). In support of hierarchy in object representation. Psychological Research, 60(3), 134-143.

Vetter, T., Hurlbert, A., & Poggio, T. (1995). View-based models of 3D object recognition: invariance to imaging transformations. Cerebral Cortex, 3, 261-269.

Vetter, T., Poggio, T., & Bulthoff, H. H. (1994). The importance of symmetry and virtual views in three-dimensional object recognition. Current Biology, 4(1), 18-23.

Wachsmuth, E., Oram, M.W., Perrett, D.I. (1994). Recognition of objects and their component parts: responses of single units in the temporal cortex of the macaque. Cerebral Cortex, 5, 509-522.

Wagner, A.D., Gabrieli, J.D.E., & Verfaellie, M. (1997). Dissociations between familiarity processes in explicit recognition and implicit perceptual memory. Journal of Experimental Psychology: Learning, Memory, and Cognition, 23(2), 305-323.

Walsh, V. (1996). Reflections on mirror images. Current Biology, 6(9), 1079-1081.

Whitlow, S.D., Althoff, R.R., & Cohen, N.J. (1995). Deficit in relational (declarative) memory in amnesia. Society for Neurosciences Abstracts, 754.

Zangemeister, W.H., Sherman, K., & Stark, L. (1995). Evidence for a global scanpath strategy in viewing abstract compared with realistic images. Neuropsychologica, 33(8), 1009-1025.

Zimmer, H. D. (1995). Size and orientation of objects in explicit and implicit memory: a reversal of the dissociation between perceptual similiarity and type of test. Psychological Research, 57, 260-273.

Appendix A: Principal Components Analysis of Eye Movement Variables

Because there is a great deal of similarity, both conceptually and statistically, between many of the 11 eye movement variables (particularly the transition variables), and because this similarity translated into a great deal of instability as to which eye movement variables would be the first to enter a step-wise discriminant or regression analysis, the correlations between all of the eye movement variables were examined. It should be noted that although the analyses reported in this section were performed in response to results reported later in Appendix B and C, they are included here since some of them become relevant to the ANOVA analyses performed in Experiment 2.

Table A.1 lists the correlations between variables using the entire data set, while Table A.2 lists the correlations using only the 25 condition by block means. Since many of the variables were highly correlated with one another, we performed a principal components analysis across all trials on all of the eye movement variables, leaving out leftfix as it was not highly correlated with any other variables. We used a principal components analysis with varimax rotation, extracting factors with eigenvalues greater than 1. Tables A.3 and A.4 show the Factor Matrix and Rotated Factor Matrix, respectively, with the loadings for the 3 factors that emerged. The first factor loads high on variables associated with the path or the fixation clusters; the second factor loads high on variables associated with the fixations; and the third factor loads high on variables associated with returning to the original fixation location. To determine how stable these factors are, we performed this same analysis by condition, by block, and by subject and found that the same three factors emerged for every condition, every block, and 37 out of 45 subjects, with 3 of the remaining 8 subjects' data showing only the minor difference of the h2t variable loading highest into the Fixation factor, rather than the Path factor. Thus we feel that these three factors represent reliable composite factors. These 3 composite variables, henceforth referred to as Path, Fixation, and Return, are included in many of the analyses discussed hereafter.

Table A.1. Correlations (r) between Variables Figured across all Data (+)

Variable	Famil	h1	h1t	h2	h2t	mfd	nclust	nfix	leftfix	retfix	rettime	s1	s1t	s2	s2t
Famil	–														
h1	-0.74	–													
h1t	-0.10	0.98	–												
h2	-0.11	0.80	0.81	–											
h2t	-0.14	0.78	0.81	0.98	–										
mfd	0.05	-0.25	-0.26	-0.39	-0.38	–									
nclust	-0.12	0.93	0.93	0.85	0.85	-0.34	–								
nfix	-0.17	0.37	0.41	0.57	0.62	-0.57	0.60	–							
leftfix	0.05	0.01	0.01	0.02	0.02	-0.03	0.03	0.02	–						
retfix	0.02	0.33	0.31	0.04	0.05	-0.15	0.30	0.17	0.05	–					
rettime	0.09	0.08	0.06	-0.28	-0.29	0.23	-0.04	-0.32	0.04	0.75	–				
s1	-0.02	0.92	0.88	0.75	0.71	-0.24	0.80	0.24	0.03	0.24	0.05	–			
s1t	-0.02	0.90	0.89	0.75	0.72	-0.24	0.78	0.24	0.02	0.22	0.04	0.98	–		
s2	-0.03	0.63	0.62	0.83	0.79	-0.34	0.62	0.30	0.03	-0.10	-0.33	0.75	0.74	–	
s2t	-0.02	0.62	0.61	0.82	0.78	-0.33	0.60	0.29	0.03	-0.10	-0.33	0.73	0.74	0.99	–

(+) Calculations used values from every trial (225) from every subject (45)

Table A.2. Correlations (r) between Variables Figured across All Condition and Block Means (+)

	Famil	h1	h1t	h2	h2t	mfd	nclust	nfix	leftfix	retfix	rettime	s1	s1t	s2	s2t
Famil	–														
h1	-0.90	–													
h1t	-0.92	0.99	–												
h2	-0.85	0.82	0.89	–											
h2t	-0.85	0.81	0.88	1.00	–										
mfd	0.63	-0.59	-0.66	0.83	-0.84	–									
nclust	-0.91	0.93	0.97	0.96	0.96	-0.78	–								
nfix	-0.77	0.69	0.78	0.96	0.97	-0.90	0.90	–							
leftfix	0.61	-0.52	-0.57	-0.52	-0.54	0.53	-0.56	-0.52	–						
retfix	0.21	-0.06	-0.20	-0.51	-0.54	0.57	-0.35	-0.60	0.31	–					
rettime	0.61	-0.48	-0.60	-0.85	-0.87	0.87	-0.75	-0.92	0.53	0.83	–				
s1	-0.68	0.87	0.80	0.49	0.46	-0.32	0.66	0.33	-0.30	0.30	-0.09	–			
s1t	-0.71	0.88	0.83	0.52	0.49	-0.32	0.69	0.36	-0.39	0.23	-0.15	0.98	–		
s2	-0.75	0.75	0.80	0.89	0.87	-0.78	0.86	0.83	-0.41	-0.49	-0.78	0.53	0.55	–	
s2t	-0.76	0.75	0.81	0.90	0.89	-0.80	0.86	0.85	-0.47	-0.53	-0.81	0.52	0.55	0.99	–

(+) Calculations used only the 25 between-subjects means

Table A.3. Factor Matrix

Variable	Factor 1	Factor 2	Factor3
h2	0.94	-0.18	
h2t	0.93	-0.18	
h1	0.92	0.29	
nclust	0.92	0.16	-0.15
h1t	0.92	0.27	
s1	0.90	0.23	0.26
s1t	0.90	0.20	0.26
s2	0.84	-0.30	0.25
s2t	0.83	-0.30	0.26
rettime	-0.15	0.93	
retfix	0.18	0.84	-0.38
nfix	0.53	-0.23	-0.73
mfd	-0.42	0.23	0.65

Note: Values less than 0.10 omitted

Table A.4. Rotated Factor Matrix

Variable	Factor 1	Factor 2	Factor 3
h2	0.95		0.14
h2t	0.95		0.12
h1	0.92	0.19	0.25
nclust	0.90	0.23	0.24
h1t	0.85	0.41	-0.17
s1	0.84	0.14	-0.36
s1t	0.83	0.14	-0.37
s2	0.82	0.43	0.18
s2t	0.82	0.46	-0.16
rettime	0.22	0.90	
retfix	-0.15	-0.79	
nfix	0.13	0.20	0.91
mfd		-0.31	0.89

Note: Values less than 0.10 omitted

Appendix B: Regression Analyses

Each of the eye movement variables were used in simple (Experiments 1 and 2) and step-wise (Experiment 2) regression analyses with familiarity to determine the relationships between these measures. Details of these procedures are discussed below.

Experiment 1

For each of the 45 subjects, the mean values of each of the 11 eye movement variables were determined for each of the 5 conditions; then, for each variable, the means for each of the 5 conditions were determined by averaging across all 45 subjects. A simple regression analysis was performed for each variable, using the five familiarity values as the regressor for each analysis. Table B.1 lists the results from this analysis. Note that some of the variables showed a strong correlation with familiarity ratings. Such a result is supportive of the use of eye movement monitoring as an indirect measure of memory.

Experiment 2

Linear. As in Experiment 1, for each of the 45 subjects, the mean values of each of these 18 measures were determined for each of the 5 conditions in each of the 5 blocks; then, for each variable, each of the 25 block by condition means were determined by averaging across all 45 subjects. A simple regression analysis was performed for each variable, using the five familiarity values as the regressor for each analysis.

R squares, mean squares, standard error of Beta, Beta, and significance levels for each of the variables figured by using the 25 block by condition means are listed in Table B.2. The r squared values and significance levels for each of the variables for each block and condition are listed in Table B.3. These simple regressions, as in Experiment 1, provide support for the use of eye movement monitoring as an effective indirect measure of memory.

Stepwise. In order to determine which variables best predicted familiarity for scenes, step-wise regression analyses were performed as well, both within block across conditions, and within conditions across blocks. Tables B.4 and B.5 list the results from the step-wise regression analyses performed across subjects by block and by condition, respectively. For the analyses performed by block across all subjects, nclust seemed to be the most effective, while the analyses performed by condition across all subjects reveals an advantage for h1 and h1t. Since there is such a great deal of variability between the eye movements of different subjects, it is more appropriate to perform this regression analysis for each subject separately. Table B.6 lists the variables and composite variables used at each step and the resultant r square values on stepwise regressions between familiarity and eye movement variables performed for each subject. For some subjects, blank spaces are present to denote that the regression analysis did not yield results that reached significance (p<.05). These results demonstrate that a significant amount of variability exists between subjects concerning which variables or composite variables correspond most with familiarity ratings.

Table B.7 provides a comparison between the two experiments of the simple regression analyses performed in block 5 between familiarity and eye movement variables. Note that while r square values tend to be higher and means square values tend to be lower in Experiment 2, variables such as mfd that figured strongly in Experiment 1 have an extremely low r square value in Experiment 2.

Table B.1. Simple Linear Regressions between Familiarity (regressor) and Eye Movement Variables (dependents) in Block 5 of Experiment 1

Variable	r squared	MS	SE B	Beta
h1	0.82*	0.128	0.024	-0.09
h1t	0.83*	0.152	0.025	-0.10
h2	0.77*	0.132	0.028	-0.09
h2t	0.77	0.163	0.032	-0.10
mfd	0.93**	1282	1.347	8.79
nclust	0.82*	2.216	0.10	-0.37
nfix	0.87*	2.148	0.079	-0.36
leftfix	0.72	0.001	0.003	0.01
retfix	0.21	0.028	0.046	-0.04
rettime	0.50	44834	30.29	51.95
s1	0.86*	0.004	0.004	-0.02
s1t	0.88*	0.004	0.003	-0.02
s2	0.86*	0.001	0.002	-0.01
s2t	0.84*	0.001	0.002	-0.01

$^*p<.05$. $^{**}p<.01$.

Table B.2. Simple Linear Regressions between Familiarity (regressor) and Eye Movement Variables (dependents) Across All 5 Blocks of Experiment 2

Variable	r squared	MS	SE B	Beta
h1	0.81**	0.184	0.004	-0.38
h1t	0.85**	0.272	0.004	-0.05
h2	0.85**	0.426	0.007	-0.06
h2t	0.71**	0.594	0.009	-0.07
mfd	0.40**	3894	1.386	5.45
nclust	0.83**	4.322	0.02	-0.18
nfix	0.60**	6.839	0.039	-0.23
leftfix	0.37**	0.003	0.001	0.01
retfix	0.04	0.071	0.023	0.02
rettime	0.37**	379210	14.7	53.82
s1	0.47**	0.003	0.001	-0.01
s1t	0.50**	0.003	0.001	-0.01
s2	0.57**	0.003	0.001	0.00
s2t	0.58**	0.002	0.001	0.00
Path	0.73**	0.101	0.003	-0.03
Fixation	0.56**	0.551	0.012	-0.07
Return	0.11	0.035	0.01	0.02

*p<.05. **p<.01.

Table B.3. Simple Linear Regressions between Familiarity (regressor) and Eye Movement Variables (dependents) by Block and by Condition in Experiment 2

Variable	r square									
	Block 1	Block 2	Block 3	Block 4	Block 5	S-S	S-D	D-D	D-S	C
h1	0.22	0.84*	0.73	0.84*	0.95**	0.87*	0.97**	0.91**	0.83*	0.24
h1t	0.24	0.85*	0.81*	0.84*	0.96**	0.92*	0.95**	0.87*	0.93**	0.38
h2	0.20	0.77*	0.78*	0.61	0.95**	0.74	0.67	0.61	0.78*	0.42
h2t	0.21	0.73	0.81*	0.63	0.92**	0.74	0.65	0.60	0.73	0.42
mfd	0.17	0.51	0.47	0.12	0.56	0.42	0.49	0.27	0.06	0.01
nclust	0.43	0.79*	0.93**	0.74	0.97**	0.89*	0.83*	0.88*	0.90**	0.21
nfix	0.10	0.64	0.76	0.55	0.81*	0.61	0.48	0.48	0.49	0.15
leftfix	0.00	0.00	0.17	0.23	0.52	0.53	0.93**	0.36	0.61	0.30
retfix	0.05	0.55	0.04	0.20	0.21	0.15	0.35	0.22	0.04	0.37
rettime	0.01	0.72	0.41	0.27	0.50	0.51	0.47	0.25	0.02	0.01
s1	0.16	0.90*	0.34	0.01	0.62	0.19	0.81*	0.74	0.41	0.25
s1t	0.27	0.85*	0.30	0.07	0.72	0.23	0.77*	0.68	0.58	0.40
s2	0.01	0.75	0.45	0.17	0.94**	0.55	0.82*	0.32	0.72	0.14
s2t	0.00	0.78*	0.52	0.16	0.90*	0.61	0.80*	0.53	0.55	0.15

S=Same, D=Different, C=Control

*p<.05. **p<.01.

Table B.4. Variables Entered in Stepwise Regressions by Block Across Subjects

Step #	Block			
	2	3	4	5
1	s1 (.90)	nclust (.93)	h1t (.84)	nclust (.97)

Note: All reached criteria (.05) in Step 1

Table B.5. Variables Entered in Stepwise Regressions by Condition Across Subjects

Step #	Condition			
	Same-same	Same-different	Different-different	Different-same
1	h1t (.92)	h1 (.97)	h1 (.91)	h1t (.93)
2	leftfix (.99)			

Table B.6. R Squares for Variables and Composite Factors in Stepwise Regressions with Familiarity Ratings in Experiment 2

Subject	r square	Variables	r square	Factors
1	.06	nclust		
2	.19	nclust/nfix/s1t	.14	fixation
3	.12	nclust/rettime/s2t	.04	fixation
7	.14	h1t/mfd	.05	path
8	.15	nfix	.13	fixation
11	.06	rettime	.08	return/fixation
12	.42	h1/h1t/nclust	.24	fixation/path
14	.10	h1t/s1		
15	.22	h2t/s2t	.08	fixation
17	.10	nfix	.11	fixation
19	.19	h1t/s1	.15	fixation/path
22	.12	h2/s2	.08	fixation
23	.13	nclust/s1	.06	fixation
24	.08	nfix	.06	fixation
25	.15	h2t/rettime/s1/s2t		
26	.03	retfix		
28	.07	h2t	.04	path
29	.04	s2		
30	.12	nfix	.08	fixation
31	.11	nfix	.09	fixation
32	.10	h1t/s1		
33	.05	s1t	.03	path
34	.11	nfix/s1t	.04	return
37	.12	nfix/rettime	.12	return/fixation
40	.04	rettime		
41	.07	retfix	.11	return/fixation
42	.12	nfix	.12	fixation/path
43	.13	nfix	.11	fixation
44	.13	nfix/s2t	.09	fixation
45	.05	s1	.05	path

Note. Analyses for subjects in which .05 limits reached before variable was entered are not listed. Variables listed in order entered.

Table B.7. Simple Linear Regressions between Familiarity (regressor) and Eye Movement Variables (dependents) in Block 5 of Experiments 1 and 2

Variable	Experiment 1				Experiment 2		
	r squared	MS	SE B	Beta	r squared	MS	SE B
h1	0.82*	0.128	0.024	-0.09	0.95**	0.042	0.006
h1t	0.83*	0.152	0.025	-0.10	0.96**	0.060	0.006
h2	0.77*	0.132	0.028	-0.09	0.95**	0.107	0.009
h2t	0.77	0.163	0.032	-0.10	0.92**	0.146	0.014
mfd	0.93**	1282	1.347	8.79	0.56	1808	4.88
nclust	0.82*	2.216	0.100	-0.37	0.97**	0.957	0.022
nfix	0.87*	2.148	0.079	-0.36	0.81*	1.98	0.087
leftfix	0.72	0.001	0.003	0.01	0.52	0.001	0.003
retfix	0.21	0.028	0.046	-0.04	0.21	0.064	0.064
rettime	0.50	44834	30.29	51.95	0.50	84710	37.8
s1	0.86*	0.004	0.004	-0.02	0.62	0.001	0.003
s1t	0.88*	0.004	0.003	-0.02	0.72	0.001	0.001
s2	0.86*	0.001	0.002	-0.01	0.94**	0.001	0.001
s2t	0.84*	0.001	0.002	-0.01	0.90*	0.001	0.001

*p<.05. **p<.01.

Appendix C: Discriminant Analyses

Discriminant analyses were also performed in both experiments in order to assess whether eye movement variables alone could be used to determine which scenes a subject was familiarity with or had prior exposure to. Discriminant analysis is a procedure for using dependent variables to separate trials into different groups. To separate trials into two groups, it uses several dependent variables and derives a weighted function that will maximize the “distance” (in mathematical space) between the trials, separating them in a manner that best corresponds to the actual grouping of these variables. This function is referred to as the discriminant function.

The discriminant function is used to compute a classification score by a procedure called jackknife validation. In jacknife validation a trial is classified using the discriminant function that was computed using all but the present trial in question. In other words, the function used to classify a trial is derived from the remaining trials in the two categories of observations (such as same-same and control). As analyses progress from trial to trial, the variables that are useful in the analysis are retained and given stronger weight in a final discriminant function that is used to classify all of the observations in the data set into two groups. From this final analysis, a classification score, which is the percentage of trials that were classified into the group that the trial actually belongs, is derived for each subject.

As the number of dependent variables that the program utilizes increases, so does the value that represents a chance discrimination rate; in other words, chance performance may be higher than 50%. For this reason, a permutation test is performed in order to determine the significance of the discrimination rate on a particular data set. A p value is calculated by randomizing the category assignment of stimuli a thousand times, redoing the discriminant analyses each time, and comparing the actual classification rate with the population of randomly arrived-at classification rates.

Four types of discriminant analyses were performed: stepwise, optimal variable selection, fixed variable, and predictive. The first three methods were used in an attempt to find a method that was most effective in discriminating between the various conditions in this experiment; the fourth method was used to examine the effectiveness of a discriminant function when its applied to trials that represent different levels of scene exposure and from which the function was not derived. Each of these methods are discussed in greater detail in the sections where they are utilized below.

Experiment 1

Step-wise discriminant analyses were performed to determine whether eye movements alone could be used to determine with which scenes subjects are familiar. Two stepwise discriminant analyses were performed. The first attempted to discriminate between eye movements in the same-same condition from those in the control condition while the second analysis attempted to discriminate between the same-different condition and the control condition. These conditions were chosen in order to determine whether the differences in familiarity between these two conditions would translate into different classification scores. The analyses performed used either a pair of variables that was set for all subjects (3 analyses), or by using an optimizing procedure in which the two variables that are optimal for maximizing classification accuracy for each subject are selected for that subject. For the 3 analyses in which the variables were set across all subjects, the variables chosen were those with the highest r square values from the regression analysis, as well as some random trial and error. Table C.1 summarizes the results of these analyses.

As the discrimination scores fall primarily around 90%, these results provide additional support for the use of eye movements as an indirect measure of memory. That classification scores using the same-same condition were not any better than the scores in the same-different condition suggests that either eye movements are not very sensitive to familiarity differences or that the nature of discriminant analysis procedures occludes such differences.

Experiment 2

To determine whether eye movement data alone could be used to sort familiar and unfamiliar trials into their appropriate groups, we performed several discriminant analyses. Four different sets of these analyses were performed and are described below.

In the first set of analyses, stepwise discriminant analyses were performed in which an unrestricted number of eye movement variables were entered into the analysis to maximize the “fit” of the discriminant function. What this means is that for each subject, different variables would be entered into the final discriminant function and that no limits were set on the number of variables that could be entered into the discriminant function; that number of variables entered were the number of variables that were needed to maximize the fit of the function. In some of these analyses, the familiarity variable was entered in order to determine whether eye movements were providing information about prior exposure in addition to what familiarity ratings already provided.

In the second set of analyses, an optimal variable procedure was performed that determined which combinations of only 2 variables would yield the highest classification scores. In this case, like the step-wise procedure, different subjects could have different variables entered into the discriminant function. Unlike the step-wise analysis, however, only two variables could be entered into the function, and these two variables would be those which together maximized the fit of the function, not necessarily the first two variables that would have been entered into the step-wise analysis for that subject.

In the third set, fixed variable discriminant analyses were performed in which the same one or two variables, chosen on the basis of previous ANOVA or regression analyses, were entered into the analysis. In these analyses, the same variables were entered into the discriminant function for all of the subjects; the variables entered were chosen on the basis of which variables which had the highest F value in the ANOVA analyses or the highest R square value in the regression analyses.

In the fourth set (Predictive), we attempted to determine the generalizability of the discriminant functions by taking equations found to be optimal in one analysis and applying them to another analysis.

It should be noted that in any of the analyses that involved the control condition, which had only 5 trials in it, the second condition entered into these analyses also had to have only 5 trials in it. Therefore, rather than using the same or different conditions, we used the sub-conditions from which these main conditions had been derived (same-same, same-different, etc.). These sub-conditions were used in all of the analyses except those in which the control condition or block 5 data were not entered.

Step-wise. In this first set of analyses, the maximum “fit” of the function was arrived at by entering as many variables into the function as necessary. Some of the analyses discriminated between conditions within each block while others discriminated between blocks within a condition. Note that the between-conditions analysis discriminates between sets of eye movement variables that arose from observing different sets of pictures, while the between-blocks analysis discriminates between eye movement variables that arose from observing the same set of pictures at different levels of exposure. In the between-conditions analysis, the control condition (unfamiliar) was discriminated from the same-same conditions (familiar) for all 5 blocks. In the between blocks analysis, block 1 (unfamiliar) was discriminated against blocks 2, 3 and 4 (familiar), for the same, different, and control conditions.

The mean of subjects’ classification scores discriminating same-same from control in all 5 blocks with either the 11 eye movement variables (Variables) or the 3 composite variables (Factors) entered into the analysis are listed in the top part of Table C.2. Only the classification scores that reached significance were figured into the mean score; the number reaching significance is listed in parentheses. The fact that classification scores in block 1, where no differences in exposure existed, were virtually identical to those in other blocks, where differences did exist according to the ANOVA analyses, makes these results difficult to interpret. The number of scores reaching significance in block 1 is generally less than those reaching significance in later blocks, however, as one would expect.

The mean of subjects’ classification scores using the 11 eye movement variables (no composite variables) to discriminate between blocks within each of the 3 conditions (same, different, and control) are listed in the top half of Table C.3. Means were determined by using only the subjects’ classification scores that reached significance and the number these reaching significance are listed in parentheses. The number of classification scores reaching significance in the same and different conditions rises across exposures, as one would expect. However, mean classification scores were higher discriminating between control images in different blocks than they were discriminating between same images in different blocks or different images in different blocks. Such a result is difficult to interpret. Since the ANOVA analyses previously demonstrated that there are differences between the eye movements for familiar and unfamiliar pictures, this difference should be revealed in discriminant analyses between conditions within blocks or between blocks within conditions or both.

In order to determine whether the eye movement variables provide any more information about prior exposure than the familiarity ratings do, step-wise discriminant analyses were performed in which the familiarity variable was entered into each analysis. Since we only wanted to determine whether the classification functions change across exposures (time), only between-condition analyses were performed. Table C.2 provides the means of subjects' classification scores using variables and composite variables (Factors). Table C.4 summarizes the frequency of which each variable was entered at each step in the analyses and Table C.5 provides the same information for the composite variables. It should first be noted that, again, classification scores in block 1 were the same as those in later blocks, no matter what variables (eye movements, familiarity, composite variables) were being used. The number of scores reaching significance, however, is less in block 1 than other blocks.

It is noteworthy that an eye movement variable was occasionally entered first into the analysis before familiarity and that in about half of the analyses, an eye movement variable was entered in step two of the analysis after familiarity. Nevertheless, little can be said about this other than that nfix, mfd, and the composite variable Fixation seem to be the most effective variables overall. Nevertheless, such results should be considered cautiously given the uninformativeness of the classification scores, as noted above.

<u>Optimal variable selection.</u> Since the step-wise procedure provided uninformative results, we tried using an optimal variable selection procedure in which only 2 variables were entered into the discriminant function. The variables were selected in such as way so as to maximize the "fit" of the function. As in the step-wise procedure, some of the analyses discriminated between conditions within each block while others discriminated between blocks within a condition. Composite variables were not used in these analyses.

In the between-condition analysis, for blocks 1 through 5, the overall classification rates for discriminating between same-different and control and for discriminating different-different and control for each block using the optimal variable procedure are listed in Table C.2. A repeated measures (block by analyses) ANOVA was performed on the classification rates for both analyses above for all subjects to determine if classification rate was effected by the condition used to discriminate against the control condition. Only a main effect of block was found [F(4, 160)=5.1, p<.001]. Individual contrast analyses were performed and found that all significant differences were found within the same-different vs. control discrimination, between block 4 and all the other blocks. Thus, the insensitivity of the between-condition discriminant analyses found in Experiment 1 was found here as well; classification score is not affected by which condition is chosen to be discriminated against the control condition, even though between-condition differences revealed by the ANOVA analyses performed on eye movements and familiarity ratings suggest that it might. Furthermore, and again, classification scores in block 1 were just as high as they were in other blocks, thereby invalidating the meaning of these scores.

Table C.6 lists the frequency of each variable's occurrence in the optimal variable procedure discriminating between conditions. As in the step-wise regression analyses, few consistencies can be seen across blocks or analyses. The variables considered optimal in one block are not in the next, and what is most effective in one analysis is not necessarily most effective in the other.

For the between-block analysis using the same-different condition and the different-different condition, the overall classification rates for discriminating between Block 1 and Blocks 2, 3, 4, and 5 are listed in the bottom half of Table C.3. A repeated measures ANOVA was performed on the classification rates for both analyses above for all subjects. Only a main effect of block was nearly found [F(3, 120)=2.5, p<.068]. Again, in contrast to what one would expect if discriminability was related to familiarity, condition and block had no effect on one's ability to discriminate between sets of eye movements.

Fixed variables. The step-wise and optimal variable procedures produced results that were uninformative. This is not likely due to the eye movement data themselves given the results from the previous ANOVA analyses but rather the over effectiveness of the procedure in deriving a function that can discriminant between two groups. Because of this, we tried using variables that were fixed for all subjects in all analyses in hopes of obtaining more informative results. In addition, we ran permutation tests in order to determine the significance of the results.

Analyses were performed in which the same one or two variables used in the analysis were fixed across subjects. The variables of nfix and h2t were used together and nfix separately since they had the highest F values in the ANOVA tables. The Fixation factor was used as it had the highest F value of the composite variables. Nclust and h1t were used together because they had the highest r square values across blocks in the simple regression analyses. No between-block analyses were performed using fixed variables.

The bottom portion of Table C.2 lists the mean of subjects' classification scores for each of the variables examined, discriminating between same-same and control in all 5 blocks. Nfix and h2t together yielded the highest classification scores, yet block 1 again had scores equal to that which emerged in later blocks, thus invalidating the meaning of these results. Thus, as with the stepwise and optimal variable procedures, the fixed variable procedure is also too rigorous. Permutation tests revealed that very few of the subjects' classification scores were statistically significant (p<.05).

Predictive discriminant analysis. Previous discriminant analyses yielded results in which classification rates between groups that were experimentally identical (all conditions within block 1 and all blocks within the control condition) were higher than classification rates between groups which were both theoretically and statistically different (on the basis of ANOVAs). In addition, the variables selected in the optimal variable selection procedure showed little to no stability across exposures and analyses. Thus, we attempted to formally demonstrate that the final discriminant function derived for one analysis was not generalizable for use in another analysis.

A predictive discriminant analysis was performed in which the equations that best discriminated between conditions in one block were used to perform the discriminant analysis in other blocks. By doing this, we hoped to determine the generalizability of a discrimination function's effectiveness. This analysis was done within subjects. Using the equations derived in blocks 2 and 4, the classification rates for discriminating between control and same-same, same-different, and different-different were determined (Table C.7). It is clear from this table that the functions that are optimal for discriminating between conditions in one block are not optimal for discriminating between conditions in other blocks. This was confirmed by performing two repeated measures ANOVAs. Using the data that emerged using the block 2 equations, a repeated measures ANOVA was performed. A main effect of block was found [F(4, 160)=145, p<.0001]. A repeated measures ANOVA using the block 4 equations also yielded a main effect of block [F(4, 164)=155, p<0001].

A predictive discriminant analysis was also performed in which the equations that best discriminated between 2 blocks (within a condition) were used to discriminate between two other blocks. This analysis was done within subjects using the same or different conditions. Using the equations derived discriminating block 1 from block 2, and block 1 from block 4, the classification rates for discriminating between block 1 and blocks 2, 3, and 4 were determined (Table C.8). A repeated measures ANOVA using the block 1 versus block 2 equations yielded a main effect of block [F(3, 120)=73.5, p<0001]. Using the data that emerged using the block 1 versus block 4 equations, a repeated measures ANOVA yielded a main effect of block [F(3, 120)=146, p<.0001] and a condition by block interaction [F(3, 120)=3.2, p<.03]. Individual contrasts reveal a significant effect of condition in the block 1 versus block 2 analysis. Just as in the between-condition analyses above, the functions that are optimal for discriminating between one pair of blocks do not generalize to discriminating between other pairs.

Both experiments yielded similar results in the discriminant analyses performed on block 5. Optimal variable selection procedures, in which the best two variables for each subject are chosen, yields between-condition classification scores of approximately 92% for both experiments. Using two fixed variables across all subjects yields classification scores of approximately 70% for both experiments. Thus, background makes little difference in our ability to discriminate between familiar and unfamiliar scenes using a discriminant analysis procedure; the relevance of the above claims, however, is likely small given the problems discussed earlier surrounding the discriminant analysis results.

In contrast to the results from the regression and ANOVA analyses, the results from the discriminant analyses were uninformative. All of the methods used (step-wise, optimal variable, fixed variable) yielded classification scores in block 1 that were virtually equal to those in subsequent blocks. The predictive discriminant analyses demonstrated that the equations most effective for discriminating between 2 conditions in one block or between 2 blocks with the same picture are not very effective when applied to a different set of blocks. The most encouraging result came from discriminant analyses in which familiarity was entered into the equation. In this analysis, eye movement variables were entered first on occasion and were often entered after familiarity, demonstrating the informativeness of these eye movement measures. Clearly, however, discriminant analyses does not represent an effective analysis tool for a measure such as eye movements.

Table C.1. Mean Discrimination Scores in Experiment 1

Variables Used	SS*C	SD*C
Numfix/s1	77	68
Mfd/h1t	75	70
Mfd/s1t	74	–
Optimal	93	90

SS=Same-same, SD=Same-different, C=Control

-- Analysis not performed

Table C.2. Mean Classification Scores from Discriminant Analyses in Experiment 2 (+)

Methods, Variables Used	Block				
	1	2	3	4	5
Stepwise, Factors (++)	84(17)	86(18)	87(22)	85(25)	90(25)
Stepwise, Variables (++)	84(13)	86(9)	82(15)	80(18)	86(16)
Stepwise, Ratings and Factors (++)	85(15)	88(18)	95(41)	93(37)	96(38)
Stepwise, Ratings and Variables (++)	88(20)	91(41)	96(43)	92(40)	96(40)
Optimal, Variables (+++)	88(44)	90(44)	89(44)	94(44)	91(45)
Optimal, Variables (++++)	90(44)	90(44)	90(44)	92(44)	92(45)
Fixed, Nfix/H2t (++)	72(44)	70(44)	74(44)	73(44)	74(45)
Fixed, Nclust/H1t (+++)	67(44)	69(44)	68(44)	70(44)	70(45)
Fixed, Nfix (++)	63(43)	67(43)	70(43)	67(42)	65(45)
Fixed, Fixation factor (++)	64(44)	63(44)	67(44)	66(44)	67(45)

(+) () denotes number of subjects included.

Only scores reaching significance listed for Stepwise analyses.

(++) Discrimination between Same-same and Control

(+++) Discrimination between Same-different and Control

(++++) Discrimination between Different-different and Control

Table C.3. Mean Classification Scores for Discriminant Analyses Between Blocks in Experiment 2

Method, Condition	Block Discriminations			
	1 versus 2	1 versus 3	1 versus 4	1 versus 5
Stepwise, Same	75(15)	78(24)	79(30)	NA
Stepwise, Different	76(18)	77(20)	76(23)	NA
Stepwise, Control	88(19)	85(22)	81(23)	NA
Optimal, SD	88	90	92	92
Optimal, DD	90	92	90	92

(+) () denotes number of subjects' scores included in mean.

Stepwise analyses include only those reaching significance.

SD=Same-different, DD=Different-different

NA=Not applicable

Table C.4. Variables Used in Stepwise Discriminant Analyses with and without Familiarity Ratings

Variables	Block 1		Block 2		Block 3		Block 4		Block 5	
	WR	WOR	WR	WOR	WR	WOR	WR	WOR	WR	WOR
rating	4		37		36		31		33	
h1	2	2	0/2	0	0/1	0	0/1	2	2	2
h1t	2	3	0	2	0/1	0	2/0/1	3	0/3	2/1
h2	0	0	0	1	0/1	1	0/2/0/1	1/1	1	1
h2t	1	0	0/2	0	0/0/1	0	1/1	3/1/1	0/1	2
s1	0/0/1	0/0/1	0/0/1	2	0/1	1/1	0	0	0/0/0/2	1/1
s1t	1	0	0	1	0/1/1	1/1	0	0/0/0/1	0/1/1/0/1	0
s2	1	1	0/0/2	0	0/1/0/1	1/1	0/1/2	0/2/1	0/2/0/1	0
s2t	0/0/1	1/0/1	0/0/0/2	0	0	1/1	1/1/0/1/1	2/0/1/0/1	0/0/3/0/1	0/2/1
nclust	0	0	1/2/1	3/1	1	2	1	2/1	0/3/1	1/1
mfd	2/2	2/2	1/1	2	2/1	3	3/2	5	0/1	3/1
nfix	3	3	1/2/0/1	3/0/3	2/4	8/1	0/1/2	2/1/1	3/2	9
retfix	4/1	4/1	0/2	0	1/1/2/1	2/1	0/4	1/1/1	1/2/0/1	1/5
rett	0	1	0/1/1	2/1	0/0/1	2	1/0/0/2	4/0/0/2	0/2/3	3/1

Note: Numbers represent number of occurrences at each step (step1/step 2/step3/step 4)

WR=With rating included, WOR=Without rating included

Table C.5. Factors Used in Stepwise Discriminant Analyses with and without Familiarity Ratings

Variables	Block 1		Block 2		Block 3		Block 4		Block 5	
	WR	WOR	WR	WOR	WR	WOR	WR	WOR	WR	WOR
Rating	4		37		36		31		34	
"Path"	3/1/1	3/1/1	1/2/1	5	0/1	2	0/3/1	3	0/4/1/1	2/3
"Fixation"	4/1	4/1	1/3	3/1	4/5/1	1	4/1	9/2	4/5/1	9
"Return"	4	4	0/1	2	0/2	2	1/2	6	0/6	5

Note: Numbers represent number of occurrences at each step (step1/step 2/step3/step 4)

WR=With rating included, WOR=Without rating included

Table C.6. Instances of Variable Use in Discriminant Analyses using Two Optimal Variables

Variable	Same-different versus Control						Different-different versus Control					
	Block 1	Block 2	Block 3	Block 4	Block 5	Totals	Block 1	Block 2	Block 3	Block 4	Block 5	Totals
h1t	28	25	45	23	24	145	30	29	42	16	28	145
h1	38	29	43	29	21	160	27	30	45	23	30	155
h2t	43	30	29	36	38	176	20	34	47	26	33	160
h2	35	24	49	29	31	168	23	33	39	31	28	154
s1t	34	29	28	27	22	140	46	42	30	24	24	166
s1	31	22	30	25	23	131	27	37	36	20	26	146
s2t	27	14	36	34	23	134	25	34	37	22	14	132
s2	34	26	34	28	26	148	26	32	32	31	21	142
nfix	26	24	28	24	8	110	59	26	39	40	19	183
nclust	28	40	54	30	19	171	32	15	39	31	14	131
mfd	44	33	34	28	26	165	34	37	32	19	26	148
retfix	29	29	29	33	24	144	33	35	34	16	18	136
rettime	28	29	41	45	30	173	25	32	31	38	32	158
leftfix	43	20	41	54	24	182	20	35	27	25	24	131
Totals	476	378	530	452	346		432	454	520	364	340	

Table C.7. Classification Scores from Predictive Discriminant Analyses between Conditions

Function Optimal in	Block				
	1	2	3	4	5
Block 2 (n=41)					
SS versus C	54	90	51	54	57
SD versus C	58	90	54	55	57
DD versus C	53	90	52	51	53
Block 4 (n=42)					
SS versus C	55	55	59	93	61
SD versus C	51	55	56	94	59
DD versus C	57	54	50	92	54

SS=Same-same, SD=Same-different, DD=Different-different, C=Control

Table C.8. Classification Scores from Predictive Discriminant Analyses Between Blocks (n=41)

Function Optimal in	Block Discriminations		
	2 versus 1	3 versus 1	4 versus 1
Block 2 versus Block 1			
Same	79	57	53
Different	79	55	58
Block 4 versus Block 1			
Same	38	53	84
Different	45	53	81

Appendix D: Multiple Views of Scenes with Delayed Exposure

A large body of research using pictures and objects as stimuli has consistently found that one set of measures, most commonly referred to as “explicit,” shows perceptually specific effects while another set of measures, most commonly referred to as “implicit,” shows no perceptually specific effects. An explicit measure, such as yes/no recognition, is one that requires the subject to reflect back to the initial exposure event in order to determine whether a stimulus was experienced before or not. An implicit measure, such as picture-fragment completion, object naming, or possible-impossible object decision, does not require the subject to recall any previous exposure with the stimulus; instead, prior exposure to the stimulus is revealed by performance in the task which is different than it would have been otherwise without the prior exposure.

The use of the terms “explicit” and “implicit” grows out of research on amnesics, such as H.M., who by definition are impaired at tasks that require explicit recall of knowledge but nevertheless show normal learning in tasks that do not require explicit recall (Graf & Schacter, 1985; Schacter, 1987; Squire, 1992; Cohen & Eichenbaum, 1993). Thus amnesics show normal priming, as well as normal cognitive and motor skill learning. Research on agnosia and prosopagnosia has also used the terms “explicit” and “implicit,” as well as the similar terms of “overt” and “covert,” to refer to the types of measures for which agnosics lack and do not lack evidence of object recognition, respectively; while these patients are impaired at recognizing an object or a person, their pattern of reaction times and skin responses during various implicit/covert tasks nevertheless discriminate between familiar and unfamiliar stimuli (Bruyer, 1991; Feinberg, et al., 1995).

This dissociation of performance between “explicit/overt” and “implicit/covert” tasks is generally thought to arise from the presence of different organic systems, or representations, in the brain. Performance on explicit tasks are thought to be processed by a declarative memory system mediated by hippocampal structures, while performance on implicit measures are thought to be mediated by those cortical brain structures organized to process the information relevant to the task (Cohen & Eichenbaum, 1993). For example, it has been suggested that priming for object naming is mediated by neural processing in Inferior Temporal (IT) cortex, as electrophysiological research has found cells in IT whose patterns of sensitivity and insensitivity to stimulus changes correspond to behavioral priming effects (Miyashita, 1993).

The association of explicit/overt measures with declarative memory and hippocampal processing, and the association of implicit/covert measures with a nondeclarative system that is not mediated by hippocampal processing has contributed to a confusion of “task” with “processor,” such that it is often assumed that an “explicit” task uses only declarative processors while an “implicit” task necessarily utilizes “implicit,” non-declarative, processors. Clearly, however, this need not be the case, for not only are both systems obligatorily engaged during any task, but performance on an implicit measure may rely on explicit remembering that is mediated by the hippocampus, particularly if the subject has had a significant amount of exposure with the stimulus. For this reason, in this thesis we have used the terms “direct” and “indirect” (Richardson-Klavehn & Bjork, 1988) to refer to the nature of the task, making no a priori assumptions about what types of processors are being engaged to perform the task. In the present experiment, however, we are directly testing whether performance in our indirect measure is being mediated by “explicit/declarative” or “implicit/procedural” processors.

As stated earlier, most research on perceptual specificity for pictorial stimuli has found that explicit (or direct) measures of memory demonstrate a greater deal of perceptual specificity than do implicit (or indirect) measures. For instance, recognition memory performance, an explicit measure, is affected if a stimulus changes in size, translation, and reflection, while these changes do not eliminate priming effects in object-naming and possible-impossible decision tasks (Biederman & Cooper, 1991, 1992; Cave & Squire, 1992; Cooper, et al.,1992; Jolicoeur, 1987; Kolers, et al., 1985; Milliken & Jolicoeur, 1992; Schacter & Cooper, 1993; Schacter, et al., 1990; Schacter, et al., 1991; Zimmer, 1995). This effect, in which implicit measures are insensitive to stimulus changes while explicit measures are sensitive to stimulus changes, has been found for changes in size and orientation (Zimmer, 1995), size and reflection (Seamon, et al., 1997), color and pattern (Cave, et al., 1996), contrast and illumination (Srinivas, 1996b), and rotation (Srinivas, 1995). In addition to these studies which show sensitivity of the explicit measures and the insensitivity of the implicit measures to such changes, other research has replicated the effects of the insensitivity of priming to changes in size (Biederman & Cooper, 1992; Fiser & Biederman, 1995), translation and reflection (Biederman & Cooper, 1991), and reflection (Srinivas, 1996a).

Contrary to this pattern of results, in Experiments 1 and 2 we found that our eye movement measures, which are, by definition, an indirect measure of memory, paralled the pattern of results of our direct measure. The focus of the present experiment is to determine whether both the direct and indirect measures are being mediated by the same representations.

In order to determine whether the eye movement results in Experiment 2, wherein eye movement variables showed the same sensitivity to view change as did the direct familiarity measure, were due to the use of processes involved in explicit remembering, we created an experimental situation in the present study in which explicit recall for a scene is allowed to fade over some delay. If explicit recall fades after a delay while eye movement measures still maintain a pattern which demonstrates sensitivity to changes in view, then the eye movement results are not mediated by explicit remembering and eye movements as an indirect measure would be unique among indirect measures in its sensitivity to such changes. If, however, the eye movements do not show such sensitivity to view change after a delay or a distinction cannot be made between old and new scenes on the basis of eye movements after the delay, then indeed, the eye movement results in the above experiments were mediated by explicit remembering.

Whitlow, Althoff, and Cohen (1995) showed subjects pictures of scenes and showed the scenes immediately afterwards to half of the subjects and 2 weeks later to the other half. Old and new scenes showed different results in both the eye movement patterns and in the explicit recognition test (a yes/no recognition judgement). But after 2 weeks, the explicit knowledge for the old scenes dropped to that of new scenes while eye movement patterns for old scenes maintained their difference from that of new scenes. This demonstrated that the eye movement effects seen immediately afterwards were not mediated by explicit knowledge, as removal of explicit knowledge did not change the eye movement results.

The present study is different than the Whitlow, et al., (1995) study in the following two ways. First, this study examines the effect of view change while the Whitlow study examined the effect of changing the spatial relationships between objects within the scene. This difference could easily lead to a different pattern of results since changing spatial relations between objects arguably changes the scene itself. Second, and most importantly, subjects in the Whitlow study performed a non-memory task, answering questions about the spatial relationships between objects in the scene displayed in front of them. Recall for scenes was only assessed during a recognition block at the end of the experiment. Thus, while the Whitlow study did not include a direct memory task during eye movement recording, the present experiment does require the subjects to make a direct memory judgement while their eyes are being monitored. This difference may explain why our data seems to show the influence of more explicit components than the Whitlow study does.

The purpose of this study is to determine whether the perceptual specificity for viewpoint that the eye movement measures demonstrated in Experiment 2 were due to processess involved in direct remembering or not. If so, this is consistent with other research using tests that require direct recall. If not, then eye movement monitoring represents a type of indirect measure that differs in its perceptual specificity profile from other indirect measures.

Method

Stimuli. The stimuli used in this experiment were composed of two sets of scenes: target (which appeared in both blocks) and distractors (which appeared only in block 1). Distractor scenes included 72 scenes that had only one view associated with them and whose content was equated with that of the target items. The target scenes included the same 45 full color photos of outdoor scenes used in Experiment 1, plus 3 additional scenes of similar content for a total of 48 different scenes. As described in Experiment 1, each scene could be shown from any one of 5 different views, and each of these views were separated by 15 degrees from the adjacent view. Unlike Experiment 1, in which a scene could be shown in any view in block 1, in the present experiment all scenes are shown in the middle view (reference image) in the first block. In block 2, target items could reappear in the same view (same condition), a different view (view condition) or appear as a control item (new condition). The appearance of all possible views of each scene was counterbalanced across subjects such that across all subjects, each view of each picture showed up an equal number of times.

In block 1, 24 target items and 72 distractor items were shown in a random order for each subject. Which 24 of the possible 48 target items showed up in block 1 varied across subjects, while every subject saw all 72 distractor items in block 1. 72 distractor scenes were used because it was found in an earlier pilot study that this number of items was required to create the memory decrement over time that is necessary for this experiment.

Participants. 32 students and staff (17 females, 15 males) at the University of Illinois volunteered for this study and were paid 5 or 8 dollars for their participation. All subjects reported normal or corrected-to-normal vision.

Procedures. The procedures for this experiment are very similar to those for Experiments 3 and 4. Subjects were instructed to examine each scene and make a judgement about how confident they were (on a 1-7 scale) that they had or had not seen the general scene (scene task) or exact picture (picture task) before. Subjects were informed that scenes may appear in a different view in the second block. After a practice session of 16 pictures, during which the experimenter made certain that subjects were clear how to respond to scenes that they recognized as being rotated, the experiment began. Each trial began with a fixation point in the center of the screen, followed by the stimulus, which remained for 1.5 seconds in block 1 and 5 seconds in block 2, followed by a prompt to make a verbal confidence judgement.

To determine whether eye movement results in Experiment 2 are due to explicit remembering or not, half of the subjects ran block 2 immediately after block 1 (immediate interval group) and the other half ran block 2 three weeks after block 1 (delay interval group). During the study phase, subjects were shown 96 pictures of scenes and were asked to make a confidence judgement about each picture. During the test phase, 48 scenes were shown and subjects made confidence judgements about the scenes. Twelve scenes were in the same view as before (same); 12 were shown in a new view (view), and twelve were completely new scenes (new). Of those shown in a new view, 6 were a 15 degree change (V1) and 6 were a 30 degree change (V2), both counterbalanced for left/right rotation; 2 levels of rotation were included in order to determine how sensitive the eye movement measures are to changes in view. In order to eliminate primacy and recency effects, each block contained 2 scene images prior to and after the 96 or 48 target and distractor images. These scenes served only as primacy and recency buffers for all subjects and were not analyzed. Subjects received a short break between blocks.

Hypotheses

In this experiment we are examining how task and study-test interval interact with the perceptual specificity of direct and indirect measures of memory for viewpoint of naturalistic scenes. As Experiments 1 and 2 already found that both the direct and indirect measures demonstrated perceptual specificity for view when the study-test interval was short, we now examine whether this effect emerges for both measures when we change the study-test interval and the subject's task.

For this experiment to be relevant, we must first replicate the results we acquired in Experiments 1 and 2; thus, we need to find, for both the direct and indirect measures, a significant difference between the same condition and the different view condition for subjects in the scene-immediate group.

To determine whether the indirect measures in Experiments 1 and 2 were being affected by explicit recall, we will examine whether lengthening the study-test interval makes the results from the direct and indirect measures diverge or not. If the results diverge such that direct measures show memory decline while indirect measures show no such decline, then we can conclude that eye movement patterns in Experiments 1 through 3 were not driven entirely by explicit remembering. If both measures show a memory-decline pattern, then we conclude that eye movements were driven by explicit remembering. We expect that explicit remembering does play a role, so we expect to falsify the null hypothesis by finding a significant difference between same-immediate and same-delay for both direct and indirect measures. If we falsify the null hypothesis only with the direct measure, then we would conclude that explicit remembering does not play a major role in the eye movement results.

Finally, to examine whether our results are affected by the type of task posed to the subject and hence to assess the relevance of the Whitlow, et al., (1995) study (which used picture recognition and found a dissociation of direct and indirect measures) to Experiments 1 and 2 (which used scene recognition and found no dissociation between measures), half of the subjects are asked if they have seen the exact picture previously while the other half are asked if they have seen the general scene before. If the type of task makes a difference, then we would falsify the null hypothesis by finding a significant interaction between condition and task.

We expect that during the test phase, confidence ratings and eye movement measures in the immediate interval group will show the same results as they did in Experiment 2, wherein both measures could be used to distinguish old scenes (same and different) from new scenes (control) and for which view sensitivity (same versus different) was apparent. For the delayed condition, however, we expect familiarity ratings to be near unfamiliar levels, since this is what Whitlow, et al., (1995) found. If the eye movements parallel this result, then we can conclude that the sensitivity of these measures to changes in view are due to direct remembering, as Whitlow found. If, however, the eye movements in the delay interval group resemble the eye movements in the immediate interval group, then we can conclude that the eye movement results in Experiment 2 were not mediated by direct remembering. It is also possible that an intermediate result could arise, wherein we would conclude that direct remembering had a partial effect on eye movement patterns.

Results

The purpose of this experiment was to determine whether indirect measures of memory (eye movements) necessarily demonstrate the same exposure and perceptual specificity effects as direct recognition measures, and whether this is dependent on the type of task (scene versus picture recognition) the subject was instructed to perform. One direct measure (confidence rating) and three indirect measures (nfix, h2t, and nclust) were recorded. These three indirect measures were chosen as they yielded the highest F values in Experiment 2 analyses.

The analyses in this section include the following: repeated measures ANOVAs using the full 3 (Condition) x 2 (Task) x 2 (Interval) design, repeated measures ANOVAs using each set of the two interval groups (collapsed across task), and repeated measures ANOVAs of condition (each interval group separately), including sublevels V1 and V2.

Analysis of full design. To determine whether the between subject variables of task and interval had any effect, a repeated measures ANOVA was performed on both the direct and indirect variables, with interval and task as the between subject variables and condition as the within subject variable (see table D.1). A main effect of task was found for the direct measure but for none of the indirect measures. This effect, in which recognition performance for old items in the picture recognition task was below that in the scene recognition task, suggests that the dissociation of direct and indirect measures in the Whitlow, et al., (1995) study may have emerged from the use of a picture recognition task rather than a scene recognition task. Had Whitlow, et al., used a scene recognition task like we used in Experiments 1 and 2, their direct and indirect measures may have paralleled one another as they did in Experiments 1 and 2. No main effects of interval were found in any variables. A main effect of condition was found for the direct ratings and h2t, and marginally for nfix, but not for nclust. No Condition x Task interaction effects were found, but Condition x Interval interaction effects were found for all variables and marginally for nclust. Since no interaction was found between condition and task, the remaining analyses collapse the two task groups (scene and picture) into one group, looking only at the effects of interval and condition.

Analysis of immediate versus delay groups. The repeated measures ANOVA described above was rerun, collapsing across task, for both the direct and indirect measures. Significant main effects of condition were found for all variables except nfix, which was marginally significant. Significant Condition x Interval interaction effects were found for all variables except nclust, which was marginally significant (see table D.1). Since these interaction effects were found, the effect of condition was analyzed separately for the immediate interval and delay interval groups.

Analysis of immediate group. In Experiments 1 and 2, both direct and indirect measures revealed differences between same and new images, as well as same and view images. To determine first whether we could replicate the exposure effects (main effects of condition) that we had in Experiment 1 and 2, a repeated measures ANOVA was performed on the direct and indirect measures for the immediate group. A main effect of condition was found for the all the variables (ratings, nfix, h2t, and nclust). The lower part of table D.1 provides the ANOVA table for these variables and the left side of Figures D.1, 3.2, 3.3, and 3.4 provide graphic illustrations of these results for these variables, respectively. In order to determine which conditions differed significantly, planned comparisons were performed on each of the variables between the 3 conditions. All comparisons with the control condition yielded significant effects. Comparisons between the same and view conditions yielded a significant difference for the direct measure but not for the indirect measures, though nfix was marginal (see table D.2). These results demonstrate that we replicated the exposure effects we found in the first two experiments when comparing same and new conditions, but not view and same or new conditions. Not finding a significant difference between view and same or new may have been due to the smaller number of subjects in this analysis (n=16) than in Experiments 1 and 2 (n=45).

Analysis of delay group. To determine whether the 3 week delay had an effect on either measure, and more specifically, whether the results of the indirect measure could be decoupled from the results of the direct measure, a repeated measures ANOVA was performed on the direct and indirect measures for the delay group. For the direct measure, a main effect of condition was found, just as in the immediate group. The lower part of table D.1 provides the ANOVA table for this analysis and the right side of Figure D.1 provides a graphic illustration of the results. Planned comparisons between conditions reveal a significant difference between same and control. The comparison between same and view, however, revealed no significant difference (p=.51), in contrast to the immediate group results. These results demonstrate that while explicit recognition of scenes is intact over a three week delay, it is also impaired relative to when recognition is tested immediately. In addition, the benefit of seeing the same picture of a scene rather than a rotated version of that scene at test is only present when the test phase follows immediately after the study phase.

For the indirect measures, no effect of condition was found in any of the variables, except nfix (see table D.1 and the right side of Figures D.2, 3.3, and 3.4); planned comparisons revealed a significant difference between the same and view conditions with the nfix variable, but not between the same and control conditions, making interpretation of this result unclear. These results demonstrate that while indirect measures are effective at detecting prior exposure when the test phase immediately follows the study phase, these measures are not effective after a delay of three weeks. This loss of an effect with the eye movement measures does not coincide with the Whitlow, et al., (1995) study in which the indirect measure did not change across the delay. This difference may be a result of using a longer delay, a different task, or a smaller n.

Levels of rotation. To determine the level of specificity of these measures for changes in view, further analyses were conducted which analyzed the 15-degree rotated images (V1) and the 30-degree rotated images (V2) separately. For these analyses on the direct measure, in which a significant difference was found between the same and view conditions, we intended to determine whether a significant difference could be found between same and V1, which would be a more specific distinction than between same and view. This analysis was performed separately for both the immediate and delay groups, collapsed across task since no interaction effects with task were found in the original analysis. For the indirect measures, in which no significant differences were found between same and view, our purpose was to determine if the same condition would be significantly different than the V2 condition, which is less specific than the distinction between same and view. A repeated measures ANOVA was performed on the direct measure of the immediate group and found a main effect of condition [$F(3,15)=27.5$, $p<.0001$]. Planned comparisons revealed significant differences between same and V1, V1 and V2, and V2 and new (see left side of table D.3). The same analysis on the delay group data revealed a significant effect of condition [$F(3,15)=13.2$, $p<.0001$] and a planned comparison revealed a significant difference between V1 and V2, and V2 and new, but not between same and V1 or V2 (see right side of table D.3). Analyses of the indirect measures revealed a pattern of V1 and V2 values that were reversed from their "expected" pattern in relation to the same and new conditions, making the results of any statistical analyses difficult to interpret. Based on these analyses on the direct measure (immediate group), we conclude that rotating a scene 15 degrees changes it enough so that it is usually identified as being different but not so much that it is ever mistaken to be a new scene.

Discussion

The primary purpose of the present experiment was to determine whether the indirect eye movement measures in Experiments 1 and 2 were being driven by the same explicit processing mechanisms that were driving the direct measures in those experiments. This was investigated because the results that we obtained in Experiments 1 and 2, in which eye movement measures showed the same sort of perceptual specificity as the direct measure, were unexpected given that previous research has often found direct and indirect measures to differ in regards to perceptual specificity. To investigate this in the present study we added a 3-week delay between study and test phases in order to create a decrement in performance in the direct measure. If we could create this decrement in the performance of the direct measure, and the results from the indirect measure did not show such a change, then we could argue that the eye movement measures are not driven by the explicit processes driving the direct measure. If however, we found that the indirect measure paralleled the decrement in the direct measure, then we would most likely conclude that the same processes driving the direct measure were driving the eye movements.

In order for this experiment to have any relevance to Experiments 1 and 2, it must first replicate the results of those experiments by showing an effect of prior exposure in both the direct and indirect variables. The repeated measures ANOVAs performed on these measures for the immediate group, collapsing across the scene and picture recognition tasks, did demonstrate a replication of the results from Experiments 1 and 2. We did not replicate this effect with the view condition using the indirect measures, but this was most likely due to the fact that in this experiment, only 16 (rather than 45) subjects were used and because subjects saw the pictures only once for 1.5 seconds (rather than multiple times at 5.0 seconds each).

Given that we were able to replicate the critical results (a significant difference between same and new) from Experiments 1 and 2 in the immediate group, analysis of the delay group's results could proceed. Analysis of the direct measure demonstrated that subjects did show a memory decrement across the 3-week delay relative to subjects in the immediate condition, but this decrement was not complete as subjects could still distinguish old items from new items. In addition, the advantage in the immediate group of seeing the exact same picture at test (same) rather than a rotated version of that picture (view) was lost in the delay group wherein no distinction could be made between the same and rotated versions of an old scene. Thus, the perceptual specificity for view of the direct measure that is seen in the immediate condition is not present 3 weeks later.

Analysis of the indirect eye movement measures revealed a complete loss of the effect of a prior exposure (in both the scene task and picture task) such that no distinctions could be made between the same and new conditions. This dissociation, in which the direct measures maintained the effects of prior exposure across the delay while the indirect measures lost this effect across the delay, is not the effect that we saw in Experiments 1 and 2 where these two measures were highly correlated. We interpret this dissocation as evidence that the two measures are not being driven by the same processor. Such an interpretation is consistent with research by Althoff (1998) where he found that while both direct (explict recall) and indirect (eye movements) correlated highly with amount of exposure (similar to Experiments 1 and 2), the two measures showed no correlation with one another.

In summary, the present study found a dissociation between the direct and indirect measures. Finding a dissocation is consistent with previous research by Althoff (1998) and suggests that the two measures are driven by separate mechanisms.

Table D.1. ANOVA Table for Primary Variables

Source	df	F				
		ratings	nfix	h2t	nclust	mfd
Full Design (n=32)						
Task (T)	1	5.4 (4.9)*	0.1 (0.5)	0.2 (0.0)	0.3 (1.0)	1.3 (5790)
Interval (I)	1	2.1 (1.9)	0.3 (1.6)	1.5 (0.3)	3.3 (12.5)+	4.1 (18240)*
Subject MS	29	(0.9)	(6.5)	(0.2)	(3.8)	(4430)
Condition (C)	2	19.3 (8.5)**	2.8 (1.0)+	3.8 (0.1)*	1.2 (0.6)	1.8 (640)
C x T	2	1.9 (0.9)	0.4 (0.1)	0.2 (0.0)	0.9 (0.5)	0.5 (165)
C x I	2	5.9 (13.4)**	8.7 (2.9)**	5.3 (0.1)**	2.9 (1.5)+	4.5 (1580)**
Collapsed across Task (n=16)						
Condition, Immediate	2	49.9 (31.6)**	6.7 (2.1)**	10.6 (0.3)**	6.4 (3.6)**	4.0 (1670)*
Condition, Delay	2	20.0 (5.6)**	5.2 (1.7)**	0.2 (0.0)	0.8 (0.4)	1.3 (267)

+$p<0.08$. *$p<.05$. **$p<.01$.

Table D.2. F Values (MS) from Contrasts between Conditions for Selected Variables, Immediate Interval Group (n=16)

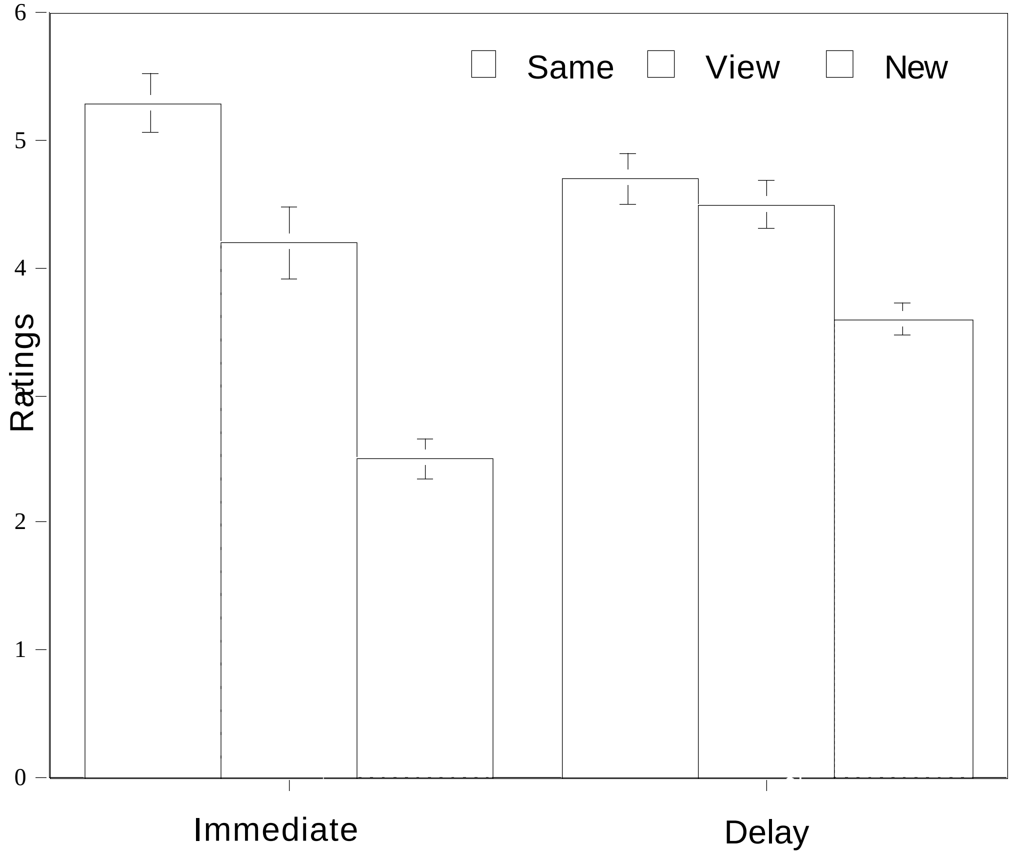
Source	F				
	ratings	nfix	h2t	nclust	mfd
Same v. Control	98.5 (62.3)**	13.3 (4.3)**	20.0 (0.5)**	12.4 (6.9)**	7.6 (3190)**
Same v. View	15.7 (9.9)**	3.4 (1.1)+	1.7 (0.1)	1.3 (0.8)	3.6 (1510)+
View v. Control	35.5 (22.5)**	3.3 (1.0)+	10.1 (0.3)**	5.6 (3.1)*	0.7 (307)

+p<0.08. *p<.05. **p<.01.

Table D.3. F Values (MS) from Contrasts between Conditions for the Direct Measure, Collapsed Across Task (n=16)

Source	Immediate	Delay
Same v. V1	5.7 (4.7)*	0.3 (0.1)
Same v. V2	21.2 (17.4)**	2.9 (1.0)
V1 v. V2	4.9 (4.0)*	4.8 (1.7)*
V2 v. New	16.9 (13.9)**	12.1 (4.2)**

*p<.05. **p<.01.



<u>Figure D.1. The Effect of Viewpoint and Study-Test Interval on Confidence of Recognition Ratings</u>

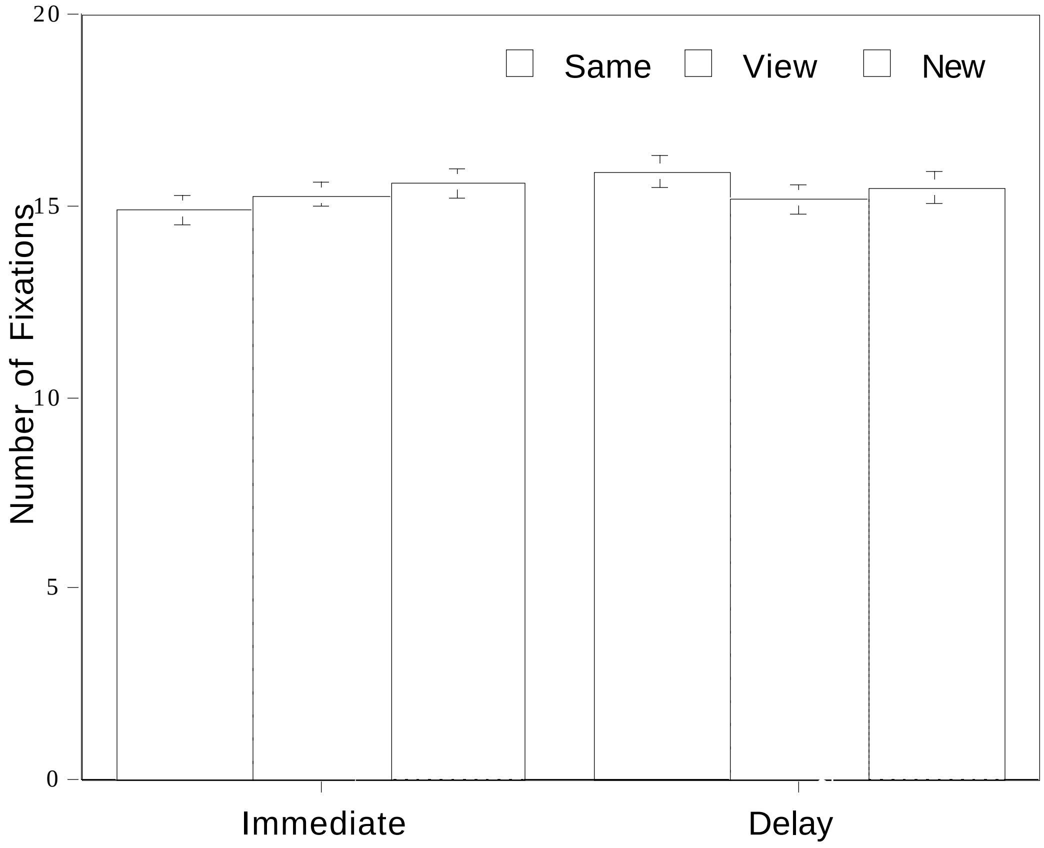


Figure D.2. The Effect of Viewpoint and Study-Test Interval on the Number of Fixations per Picture

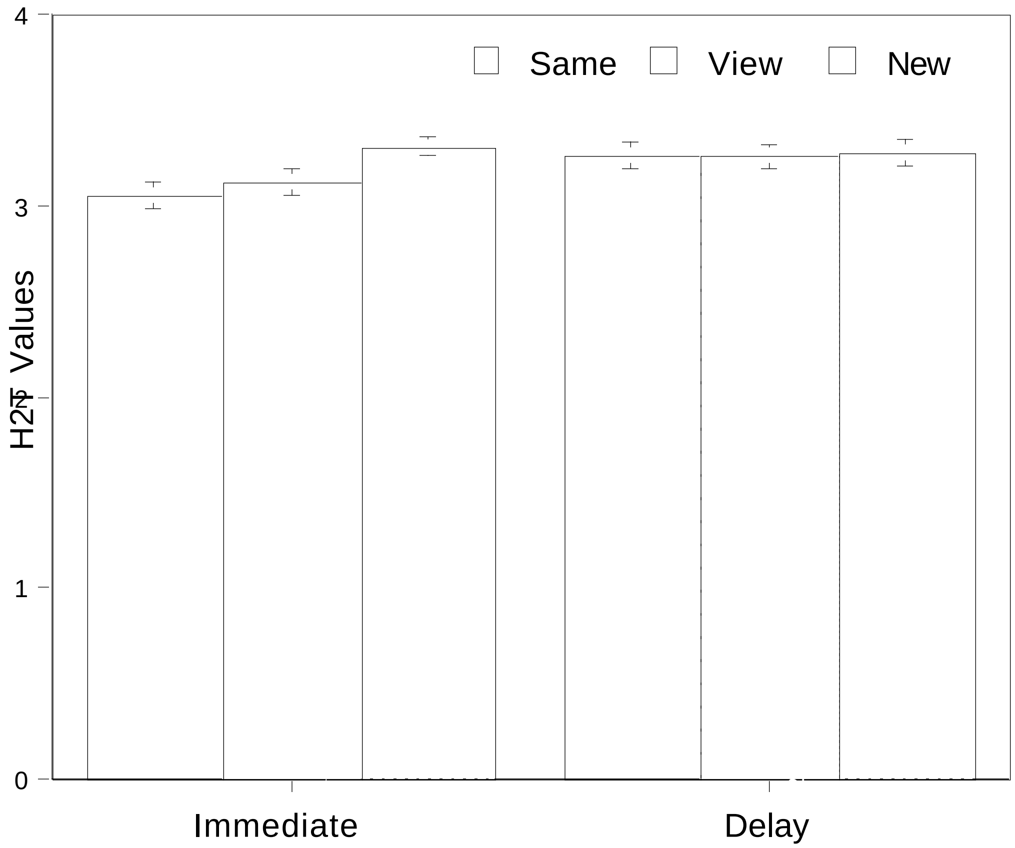


Figure D.3. The Effect of Viewpoint and Study-Test Interval on the H2T Entropy Measure

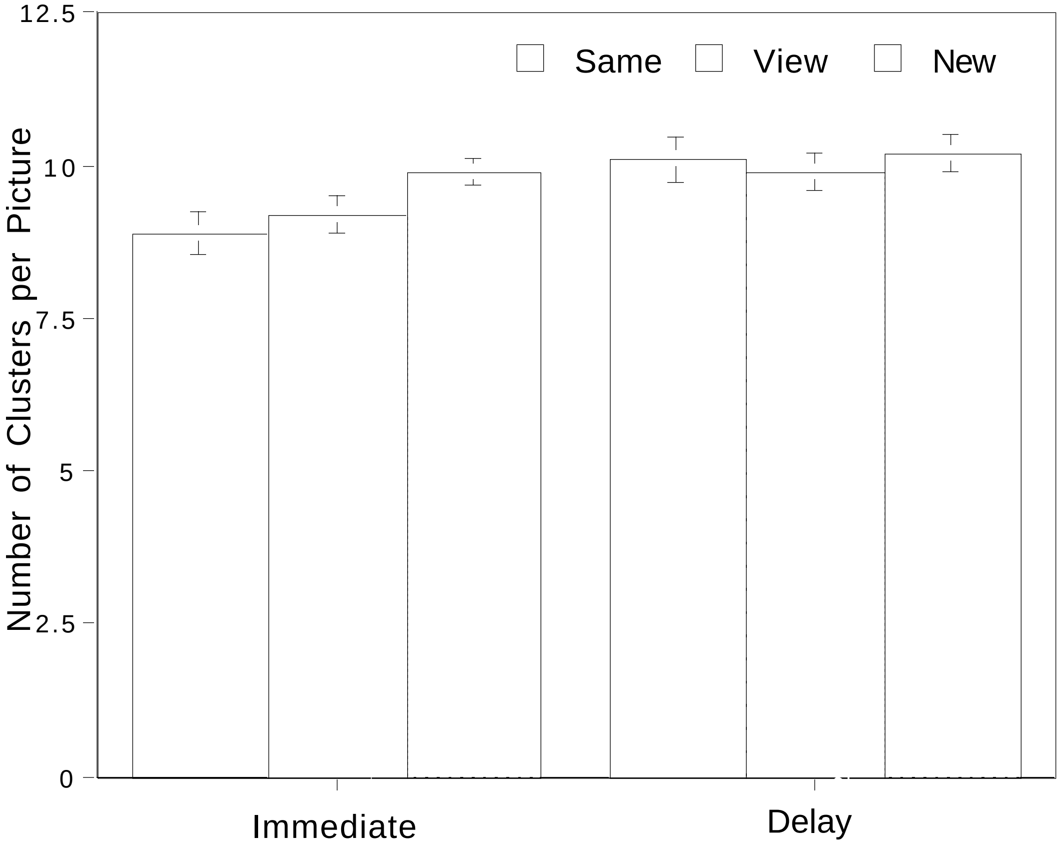


Figure D.4. The Effect of Viewpoint and Study-Test Interval on the Number of Clusters per Picture

Vita

Mark William Olson

DOB: September 1, 1968. St. Charles, Illinois.

Education

A.M. Educational Psychology. University of Illinois at Urbana-Champaign. August, 1992.

B.A. Psychology. Taylor University. Upland, IN. May, 1990.

Teaching Experience

Fall, 1997. TA for Neuroscience/Physiology 315: Structure and Function of the Nervous System. Instructor: Dr. Matilde Holzwarth.

Spring, 1997; Spring 1996; Fall, 1995. TA for Medical Neuroscience. Instructor Dr. Eve Gallman

Fall, 1990 to Fall, 1992. TA for Educational Psychology 211: Introduction to Educational Psychology. Instructor: Dr. David Zola.

Research Experience

Spring, 1997. RA for Dr. Arthur Kramer. Topic: 3-D Visual Attention.

Summer and Fall, 1996. RA for Dr. Arthur Kramer. Topic: EEG and Spatial Attention.

Spring, 1994 to Spring, 1995. RA for Dr. Neal Cohen. Topic: Eye Movement

Monitoring and Memory.

Fall, 1992 to Spring, 1993. RA for Dr. Joanne Vining. Topic: Comprehensive Review of the Literature on Environmental Values.

Papers Presented

Causation and Volition: A Perceptual Control Theory Perspective. Presented July, 1992 at the Perceptual Control Theory Conference in Durango, Colorado

Works in Progress

Olson, M.W. & Parsons, R. (in process). Eye movements predict aesthetic preference of natural scenes.