THE ROLE OF FAMILIARITY ON CHANGE PERCEPTION

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF INFORMATICS
OF
THE MIDDLE EAST TECHNICAL UNIVERSITY

BY

HACER KARACAN

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY
IN
THE DEPARTMENT OF COGNITIVE SCIENCE

JULY 2007

Approval of the Graduate School of Informatics

____

Assoc.Prof.Dr.Nazife BAYKAL
Director

I certify that this thesis satisfies all the requirements as a thesis for the degree of Master of Science/Doctor of Philosophy.

____

Prof.Dr.Deniz ZEYREK
Head of Department

This is to certify that we have read this thesis and that in our opinion it is fully adequate, in scope and quality, as a thesis for the degree of Doctor of Philosophy.

__________________________	__________________________
Prof.Dr.Hasan Gürkan TEKMAN	Assoc.Prof.Dr.Kürşat ÇAĞILTAY
Co-Supervisor	Supervisor

Examining Committee Members

Assist.Prof.Dr. Bilge SAY	(METU, COGS)	____________________
Assoc.Prof.Dr.Kürşat ÇAĞILTAY	(METU, CEIT)	____________________
Prof.Dr.Hasan Gürkan TEKMAN	(ULUDAĞ, PSY)	____________________
Assoc.Prof.Dr. Cem BOZŞAHİN	(METU, CENG)	____________________
Assist.Prof.Dr. Annette HOHENBERGER	(METU, COGS)	____________________

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this wok.

Name, Last name : Hacer Karacan

Signature : ____________________

ABSTRACT

THE ROLE OF FAMILIARITY ON CHANGE PERCEPTION

Karacan, Hacer
Ph.D., Department of Cognitive Science
Supervisor: Assoc. Prof. Dr. Kürşat Çağıltay
Co-Supervisor: Prof. Dr. Hasan Gürkan Tekman

July 2007, 162 Pages

In this study the mechanisms that control attention in natural scenes was examined. It was explored whether familiarity with the environment makes participants more sensitive to changes or novel events in the scene. Previous investigation of this issue has been based on viewing 2D pictures/images of simple objects or of natural scenes, a situation which does not accurately reflect the challenges of natural vision. In order to examine this issue, as well as the differences between 2D and 3D environments, two experiments were designed in which the general task demands could be manipulated. The results revealed that familiarity with the environment significantly increased the time spent fixating regions in the scene where a change had occurred. The results support the hypothesis that we learn the structure of natural scenes over time, and that attention is attracted by deviations from the stored scene representation. Such a mechanism would allow attention to objects or events that were not explicitly on the current cognitive agenda.

Keywords: Attentional capture, Eye movements, Scene memory, Virtual Reality, Eye Tracking.

ÖZ

AŞİNALIK HİSSİNİN GÖRSEL DEĞİŞİKLİKLERİ ALGILAMA ÜZERİNDEKİ ROLÜ

Karacan, Hacer
Doktora, Bilişsel Bilimler Bölümü
Tez Yöneticisi: Doç. Dr. Kürşat Çağıltay
Ortak Tez Yöneticisi: Prof. Dr. Hasan Gürkan Tekman

Temmuz 2007, 162 Sayfa

Bu çalışmada, görsel dikkati doğal görsel ortamlarda kontrol eden mekanizmalar incelenmiştir. Sözü edilen görsel ortamlarda gelişen aşinalık hissinin gözlemcileri ortamda oluşabilecek değişikliklere ve/ya yeni olaylara karşı daha hassas yapıp yapmadığı araştırılmıştır. Bu konunun incelendiği daha önceki çalışmalarda bulgular iki boyutlu fotoğraflar ve/ya resimler kullanılarak elde edildiği için doğal görmedeki karmaşık yapıları tam olarak yansıtamamaktadır. Doğal ortamlarda görme konusunu daha güvenilir bulgularla inceleyebilmek ve iki boyutlu ve üç boyutlu ortamlardaki görüş farklılıklarını daha iyi anlayabilmek için, çalışma kapsamında genel görsel algı gereksinimlerinin uygulanabileceği iki farklı deney tasarlanmıştır. Bu deneylerden elde edilen bulgular görsel ortama karşı gelişen aşinalık duygusunun ortamdaki bir değişikliğe odaklanma süresini anlamlı bir şekilde arttırdığını göstermektedir. Ayrıca, bu bulgular insanların doğal görsel ortamları zamanla öğrendiğini ve görsel dikkatin o anki gözlemlerinin bilgi haznelerinde tuttukları betimlemelere göre oluşturduğu farklılıklardan kaynaklandığı hipotezini desteklemektedir. Bu şekildeki bir mekanizma insanların görsel dikkatlerinin bilişsel ajandalarında bilinç üstü bir şekilde tutulmayan objeler ve/ya olaylar üzerine çekilmesine olanak sağlayabilmektedir.

Anahtar Kelimeler: Görsel Dikkat, Göz Hareketleri, Görsel Hafıza, Sanal Gerçeklik.

This work is dedicated to;

Cüneyt – The Meaning of Life and Happiness for me

&

My Parents and Brothers – The Sunshine on me

ACKNOWLEDGMENTS

I want to express my sincerest appreciation to my advisor Assoc. Prof. Dr. Kürşat Çağıltay for his kind supervision, guidance, suggestions, and friendly attitude throughout this study. I also wish to express my appreciation to my co-advisor Prof. Dr. Hasan Gürkan Tekman for his patience, valuable guidance and supervision throughout this study. I should also thank Prof. Dr. Mary Hayhoe for her precious guidance and Assoc.Prof. Dr. Cem Bozşahin, Assist. Prof. Dr. Annette Hohenberger and Assist. Prof. Dr. Bilge Say for their vauable comments and suggestions about this study.

I would also like to express my thanks to Feza Taş, Yalın Baştanlar, Cemile Serçe, Özge Demirkol, Nadide Seyhun, Çağlar Özyurt, Serkan Şedele, Aslı İnan, Tolga İnan, Nigar Köktaş, Habil Kalkan, Selim Nar, Üstün Kaya, Emre Keskin, Kelly Chajka, Jelena Jovancevic, Brian Sullivan, Keith Parkins and all of my friends at the Informatics Institute for their help, comments, encouragement through the hard times during the study, and for their friendship. I should also say that it would be impossible to finish this study without the help and support of my friends Sibel Gülnar and Pınar Onay Durdu.

Special thanks are also extended to my parents Nursel Üke, Mehmet Üke, Umur Karacan and Kemal Karacan for being by my side during my trials and tribulations and assuring me, when I was most hopeless, that good days were yet to come.

I would also express my sincerest appreciation to my brothers Barış Üke and Ömer Can Karacan for always standing by my side during this study and to my sister Ayşe Sayın for being my dreamsharer.

Finally, my endless and profound thanks and love would go to my beloved husband Cüneyt Karacan. All the beauties, values and happiness he brings to my life and the love he gives to me without any expectations make me the luckiest person in the world.

This work was supported by NIH grants EY05729, RR and TUBİTAK under grant SOBAG 104K098. I wish to acknowledge them for making this study possible by providing us with the eye-tracking systems and other necessary equipments at the University of Rochester and the Middle East Technical University.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

2D : Two Dimensional

3D : Three Dimensional

ANCOVA : Analysis of Covariance

ANOVA : Analysis of Variance

AR : Augmented Reality

AV : Augmented Virtuality

CAVE: Cave Automatic Virtual Environment

CIB : Ceiling-Hiball Interface Board

GABA : Gamma Amino Butyric Acid

HMD : Head Mounted Display

LTM : Long Term Memory

LCD : Liquid Crystal Display

METU : Middle East Technical University

MTL : Medial Temporal Lobe

PC : Personal Computer

RV : Reality - Virtuality

SGI : Silicon Graphics, Inc.

SPSS : Statistical Package for the Social Sciences

STM : Short Term Memory

VA : Visual Acuity

VE : Virtual Environment

VR : Virtual Reality

CHAPTER 1

INTRODUCTION

It is commonsense knowledge that attention is important for learning. We need to deploy our attentional resources to learn something efficiently and effectively. Similarly, we may also need to learn our surroundings in order to be able to deploy our attention. In this study, learning is defined as a product of an exposure to a natural visual stimulus and the general issue of attention and its relationship with this kind of learning process (i.e. familiarity) is investigated in virtually modeled natural scenes.

For this purpose, two Virtual Reality Environments were created and experiments that would provide the required answers were conducted in these environments. Eye-tracking technology was also used to follow the position of the observers' eye-gaze, which made it possible to measure the attended location in the environment.

The motivation behind the usage of virtual reality (VR) for cognitive purposes is twofold. Firstly, it is easy to use virtual reality applications for examining cognitive issues for some cases that are difficult to handle in real-world environments. Secondly, I wanted to provide a new perspective that the results of cognition experiments bring into the development of virtual reality technology since the future of this technology will be shaped according to the users' physical and cognitive demands as well as their performance levels in different kinds of virtual environments. In this study Immersive VR technology is used to control the experimental environment in which sudden changes occur.

1.1 Problem Statement

This study examines the differences in visual attention under different levels of familiarity in a virtually modeled natural environment. The experiments in this study investigate the following research question:

- Does familiarity with the environment make participants more sensitive to changes or novel events in the scene?

The expected results of this study shall give strong evidence for the importance of learning for allocation of attention. Since one of the experiments of this study was conducted in an immersive interactive 3D environment, this will give insights about the real-world situations. Moreover, the results will provide additional evidence to the ones that were previously presented with 2D pictures. A Desktop Virtual Reality environment was also used to understand the general issue and will give an original answer to the research question while also providing information about the differences and similarities between two kinds of virtual environments (immersive vs. desktop) for this issue.

1.2 Significance of the Study

We receive information from the outside world by using our five senses, and among these five, vision is the dominant sense for interacting with our world. In order to understand this interaction better, we need to know what controls the information we get from the environment. The issue is important, because how information from the real world is represented can greatly affect how easy is it to do different things with it (Marr, 1982).

For the information we get from the environment to be useful and detailed, we have to deploy our visual attention to the area from where the information comes. In addition, it is important to know the control mechanism of this visual attention in natural scenes.

Control of the locus of gaze and deployment of attention in natural environments is largely unexplored. On this issue, it was claimed that detailed visual information is reliably retained in memory from previously attended objects (Hollingworth & Henderson, 2004). When it comes to attending to changes in a scene, some studies showed that observers are insensitive to changes in scenes made during saccades, or when masked by some other kind of transient such as a blank screen or a blink (Rensink, 2000; Simons, 2000a; O'Regan, 1999). Even if we rely on this small amount of previous studies on this issue, this may not generalize the fact since they used either simple stimulus arrays or 2D images of naturalistic scenes. In real world cases there exist no 2D images, instead we interact with 3D environment. Therefore, unlike the previous studies, I choose to use an interactive environment including 3D images to study the general issue of what controls attention in natural scenes. For this reason, an immersive environment was used to see what happens in "real" interactive scenes.

Previous findings reveal that there is some global representation of the scene, and that as long as the new information is consistent with this, there is no need to update the representation. In this case, a changing object in a scene may only become visible if it can activate some process that reports a *mismatch*. It may be difficult to detect changes in the scenes used in change blindness experiments because the internal referent is ill-defined. This lack of internal definition results from the fact that observers are presented with unfamiliar scenes. They do not know how the scenes should look like. Thus an unexpected object might be detected more frequently in an environment that is familiar and therefore well learnt. Such a finding would go a long way to resolving some of the questions raised by change blindness like the reasons for suprising difficulty observers have in noticing large changes.

In order to examine the tradeoffs explained above, two experiments were designed in which the general task demands could be manipulated. These experiments were used to probe the question that asks whether the changes are more likely to attract attention if participants become familiar with an environment.

The results of this study will be helpful to explain the nature of change perception for four different change types. The prioritization of abruptly appearing and disappearing objects in real-world scenes was previously explored (Theeuwes, 1991; Mondy & Coltheart, 2000; Boot, Kramer & Peterson, 2005; Brockmole & Henderson, 2005). This study introduces additional data for these evidences as well as two more change types (i.e. replacement and movement) that was not investigated much before (Grimes, 1996; Franconeri & Simons, 2003). It also brings a new insight about the analysis of disappearing objects. In order to understand the detection rate of disappearing objects, previous studies compared the eye movements on the existing object and the empty space where that object was used to be. This may be misleading since comparison of an object and an empty spot is not cognitively valid. An empty spot cannot be as informative as a present object. Because of this reason, the possibility of looking at the empty spot of the disappeared object was used as the control condition in this study.

In the experiments that are mentioned above, eye tracking and virtual reality technologies were used for an interesting and meaningful research. There are some important reasons for analyzing eye movements in scene viewing. First of all, eye movements are critical for the efficient acquisition of visual information during complex visual-cognitive tasks. In addition, the data that was gathered by analyzing eye movements provide a measure, which presents the processing of visual and cognitive information in an unobtrusive way (Henderson & Hollingworth, 1999). The usages of virtual environments also provide more naturalistic results for the change perception literature while making contributions on change blindness phenomena discussed above.

1.3 Organization of Dissertation

This chapter has described the problem addressed by this thesis in the context in which it arises, as well as providing a description of its contributions. The next chapter includes the necessary definitions of terms that the study is based on.

Chapter 3 examine the relevant literature, summarizing other studies and placing the work in context. This chapter also introduces this study by spelling out its contributions on the literature.

Chapter 4 describe the behavioral experiments while giving information about the virtual environments in which the experiments of the study were conducted and emphasizing the aspects of its construction. The design of these experiments are also discussed and the information about the participants, apparatus, procedure and measurements are given in this chapter. The results of the experiments are also interpreted and discussed in this section.

Finally, Chapter 5 include a general discussion about the results of the experiments. It also includes the summary of the work and the examination of the conclusions that might be drawn and then speculate on the future of virtual reality in visual perception applications while discussing the limitations of this study.

CHAPTER 2

DEFINITION OF TERMS

In order for the concepts of this study to be more understandable some brief definitions are given in this section. These concepts include "visual attention", "change blindness", "spatial learning" and "familiarity". Furthermore, "virtual reality" and "eye-tracking" technologies are explained to clarify the procedure.

2.1 Visual Attention

Perception comprises a range of different ways of informing ourselves about the environment for a variety of different purposes including understanding, recognition and the control of action. One important tool adapting the visual system to different perceptual tasks is the set of control systems that we call attention (Treisman, 2006).

Attention is one of the most interesting aspects of cognitive psychology since it is a cognitive process, which gives rise to conscious awareness (Naish, 2005; Posner, 2003). It can be described or defined in numerous ways. For example, it can be described as the allocation of information processing resources to a specific source of information or simply the mind's ability to focus and concentrate. Furthermore, Tsotsos (2001) describes attention as a set of strategies that attempts to reduce the computational cost of the search processes inherent in perception.

In addition to this general concept there are subcomponents of attention and in this study our main concern is the visual part of it. Visual attention is one of the most important features of the human visual system and it can be described as a mechanism which filters out redundant visual information and detects the most relevant parts of our visual field (Le Meur, Le Callet, Barba & Thoreau, 2006).

In our daily lives, we have to process massive amounts of incoming visual information. On the other hand, there is a requirement for nearly real-time processing. In order to achieve this goal, we may employ a serial strategy by which an attentional spotlight rapidly selects circumscribed regions in complex visual scenes for further analysis (Itti, Goal & Koch, 2001).

The limited visual and computational resources available during the perception of a human action make visual attention mechanism essential. It can be focused narrowly on a single object or spread over several or distributed over the scene as a whole. In addition to increasing or decreasing the number of attended objects, these different deployments may have different effects on what we see (Treisman, 2006). When we observe our visual environment, we do not perceive all of its components as being equally interesting. Some objects automatically “pop-out" from their surroundings, that is, they attract our visual attention immediately (Itti, 2000).

The singular idioms describing the selective nature of attention are the “where” and “what”. The former corresponds to the visual selection of specific regions of interest from the entire visual field for detailed inspection while the latter corresponds to the detailed inspection of the spatial region through a perceptual channel limited in spatial extent. Considering visual attention in these terms, we would expect the eye movements work in a way that vision behave in a cyclical process composed of the following steps (Duchowski, 2003):

1. Given a visual stimulus, the entire scene is first seen mostly in parallel through peripheral vision (i.e. at low-resolution).

   - at this stage, interesting features may "pop out" in the field of view, in a sense directing attention to their location for further detailed inspection.

2. Attention is thus disengaged from the foveal location and the eyes are quickly repositioned to the first region which attracted attention.

3. Once the eyes complete their movement, the fovea is now directed at the region of interest and attention is then engaged to perceive the feature under inspection at high resolution.

These steps can help to understand the nature of visual attention while explaining the essential role of eye movements in this mechanism.

2.2 Change Blindness

The term 'change blindness' refers to the surprising difficulty observers have in noticing large changes to visual scenes (Rensink, O'Regan & Clark 2003; Simon & Levin, 1997). The majority of tasks used to explore change blindness involve a change over time. While perceiving a scene, a change occurs during a saccade or video cut, perception continues, and then the participant indicates in some way whether there were changes to the scene during the viewing. If the timing of the image manipulation is successfully correlated with the eye-movements of the observer, most of them miss seven out of ten changes (Grimes, 1996). This is a task with important similarities to a recognition memory task. It can be considered to be a "same or different" recognition task in which the target and distracter are alternative views of a scene perceived earlier (Pani, 2000).

Work on change blindness try to clarify the nature of the juncture between perception and cognition. Here, the main question is that how perception and cognition are related. Perception provides us with very rich information and this information is important for high fidelity sensing of the world and for a low-level organization of information that will allow flexible learning. On the other hand, much of the perceived information turns into the form of memory, knowledge and reasoning. Since these systems have a limited capacity for handling the information, this information is characterized by selection. In other words, we process two types of information in perception and in higher cognition and the interaction between these different information structures can give us some clues about the relation between perception and cognition (Pani, 2000).

Work on change blindness has also helped defining concepts which were not previously well articulated (Simons & Rensink, 2005). For example, it forced a clarification of the distinction between motion perception, change perception, and difference detection, which are the notions that often were conflated in earlier work. "Motion" started to be defined as variation referenced to location, whereas "change" is defined as variation referenced to structure. "Change" is also distinguished from “difference” such that the former refers to the transformation over time of a single structure while the latter refers to a lack of similarity in the properties of two structures. Under the light of these definitions, “Motion perception” is now more clearly understood as the detection of unorganized flow at a location, “change perception” as the detection of an ongoing transformation of a structured object, and “difference perception” as an inferential comparison of the current stimulus with traces in long-term memories (Rensink, 2002).

Change blindness has contributed to a resurgence of the study of scene perception and the dynamics that underlie it. More broadly, change detection and change perception can be considered special cases of event perception, and the concepts and techniques developed in this literature could become useful tools for understanding the perception of dynamic events more generally (Simons & Rensink, 2005).

2.2.1 In-attentional Blindness

In-attentional blindness occurs when observers do not notice an unexpected object in a display. In a typical in-attentional blindness procedure, an unexpected, task-irrelevant object appears either in the final trial of a small set of experimental trials or for some duration during a continuous task. Following the usual task–response on the critical trial, participants are asked to report whether they were aware of any extra task-irrelevant stimulus, or anything unusual on the screen[1]. Findings show that participants often fail to notice the unexpected task-irrelevant stimulus. By contrast, the same stimulus is often

[1] See Mack & Rock (1998) for detailed information about the procedure.

detected on a following control trial in which participants do not perform the task but instead pay attention to whether there is any extra stimulus on the screen (Cartwright-Finch & Lavie, 2007).
Rensink (2002) claims that the critical factor for detecting changes in a scene is attention. In other words, to see an object change, it is necessary to attend to it. He proposed the *Coherence Theory* in which he claims prior to focused attention humans produce several object representations, but they lack stability and are quickly replaced by new objects while focused attention produces a very detailed and long-lasting representation of one object. By this way only, a change will be perceived if a new stimulus replaces the object. When focused attention is removed, representation disintegrates and returns to the first stage.

As a conclusion, Rensink made the following proposal with his colleagues O'Regan and Clark (2003):

- visual perception of change in an object occurs only when that object is given focused attention;

- In the absence of focused attention, the contents of visual memory are simply overwritten (i.e. replaced) by subsequent stimuli, and so cannot be used to make comparisons.

2.3 Spatial Learning & Implicit Memory

The most commonly offered definition of learning is that the process by which relatively permanent changes occur in behavioral potential as a result of experience (Anderson, 1995). Another definition of learning is that a mechanism for adapting to novel situations (Cech, 1998). Memory, on the other hand, was defined as the relatively permanent record of the experience that underlies learning (Anderson, 1995). These definitions reveal that learning is a process of change while memory is the exclusive *product* of this *process*.

In this study, the kind of learning that is important for answering the research questions is the spatial one, since this study proposes that learning the surrounding environment is crucial for change perception. In order to understand the interactions between the process of spatial learning and the *product* of this *process*, we need to look at the role of spatial structures in abstract thought, which is an important issue for cognitive psychology. These spatial structures, also known as spatial schemas, share the general qualities of other types of schemas, which aid cognition because they are organized. They have a familiar structure, and people can rely on that structure to facilitate memory (Gattis, 2001). The difference of spatial schemas from other types of schemas is their source in abstract thought. Since they are first acquired in purely spatial context, they must be transferred to a form which is usable for memory structures. For example, many people experience the sensitive relationship between personal history and memory for place when re-visiting a place after many years. The experience of re-visiting a place points out that space can sometimes be a more powerful organizer of memory than time (Gattis, 2001).

Recent studies suggest that normal adults can rapidly acquire spatial contextual knowledge. For example, when participants search for a T target among L distractors, search speed was found to be faster for displays that have previously been seen than for novel displays (Chun & Jiang, 1998).

The distractor locations on repeated displays form a consistent visual context, which guides spatial attention to the associated target's location. This paradigm, known as *contextual cuing,* was used in order to investigate how implicit learning of regularities generated either by specific elements or by more general semantic categories in the context can facilitate visual-scene analysis. The results showed that such learning is surprisingly powerful (Jiang & Song, 2005). It occurs after just five or six repetitions and lasts for at least a week (Chun & Jiang, 2003; Jiang, Song, & Rigas, 2005). It is also implicit, because participants rarely notice the repetition, nor can they recognize repeated displays after learning.

Chun and Jiang (1998) hypothesized that people benefit from repeated displays in visual search because they have learned the configuration of repeated displays. In another study, (Jiang & Song, 2005) it was shown that spatial context learning applies not only to visual search, but also to change detection tasks. When participants have to remember a few spatial locations for a brief amount of time, their performance is enhanced on repeated displays.

In visual-scene analysis literature, many authors have agreed that visual representations are supported by a scene schema stored in Long Term Memory (Henderson & Hollingworth, 1999; Rensink, 2000). Furthermore, in order to answer the question of whether implicit learning contributes to the construction of scene schemas, Goujon, Didierjean and Marméche (2007) also used the contextual cueing paradigm developed by Chun and Jiang (1998), in an attempt to experimentally reproduce contextual-regularity learning effects. While reporting the results of this study, the authors claimed that during the analysis of an image or a visual scene, the visuo-cognitive system seems to be able to implicitly encode and store not only spatial relationships between the specific features of contextual elements, but also relationships bearing on certain categorical properties of those elements. Therefore, Goujon, Didierjean and

Marméche answered their initial question as both implicit learning and retrieval mechanisms should contribute to the construction of scene schemas.

2.4 Familiarity

> Recognition is an essential but challenging component of daily social life. By connecting the present to the past, recognition maintains people's sense of self and enables them to follow their goals by placing their immediate experience in the context of a continuing narrative with recurrent people and places (Monin, 2003, p.1035).

Recognition memory is considered to be supported by two different memory processes; a sense of *familiarity* gained from previous exposure to particular stimuli and the explicit *recollection* of information about a previous event (Mandler, 1980). The first process, familiarity, is the knowledge of having seen or experienced some stimulus before, but having little information associated with it in your memory. Recollection, on the other hand, is characterized by richer associations (for review, see Yonelinas, 2002). Relative to recollection, familiarity is claimed to be less attention demanding, more perceptual than conceptual, require less processing time, and be less likely to involve conscious awareness (Yonelinas, 2002).

In general, the term familiarity is used to describe the fact that perceivers respond differently to novel stimuli than they do to stimuli with which they have had some experience (Haber & Hershenson, 1973). This is a more logical definition than those psychological definitions which purely dissociate familiarity from other memory processes as stated above. Similarly, in this study, "familiarity" is defined as a person's sense of being familiar to a scene after his/her exposure to a visual stimulus (except for in subsections 2.4.1 &

2.4.2). Here, the degree of familiarity is not defined as a property of stimulus; rather it is a property of the observer. Since different observers may have had different amounts of experience with the same stimulus, familiarity with respect to a particular stimulus may be different for different observers.

2.4.1 The Differences between Familiarity and Implicit Memory

In order to be clear about the type of familiarity that is mentioned in this study, its differences from implicit memory are discussed in this subsection. The reason for stating the differences between these two concepts is the possibility of confusing the term "familiarity" with any kind of implicit processes in mind.

There are six main functional dissociations between familiarity and performance on perceptual implicit memory tasks, indicating that familiarity is not identical to perceptual implicit memory (Yonelinas, 2002):

1. encoding manipulations (such as comparing deep / shallow processing, generated / read items, allocation of full / divided attention) generally increase estimates of familiarity, but they do not generally influence perceptual implicit memory.

2. familiarity is greater for generated than seen/read items, whereas the opposite is true for perceptual implicit memory (i.e. perceptual implicit memory is better for seen/read items than generated ones).

3. familiarity exhibits a picture-superiority effect, which reflects the fact that memory for words is better if the item was studied as a picture than if it was studied as a word. On the other hand, perceptual implicit

memory shows the opposite effect (Wagner, Gabrieli, & Verfaelie, 1997).

4. benzodiazepines[2] generally reduce familiarity to some extent, whereas these drugs typically do not influence perceptual implicit memory.

5. familiarity is reduced in amnesic groups with extensive temporal lobe damage, and these groups typically do not exhibit deficits on perceptual implicit memory tests (Gabrieli, 1998).

6. the neuroimaging results from perceptual priming studies often implicate extra striate regions in the occipital lobe (Cabeza & Nyberg, 2000) and there is little evidence linking this region to familiarity.

We can also talk of another hypothesis about the dissociation if we think familiarity as implicit learning which is said to occur when there is an increase in task performance without an accompanying increase in verbal knowledge about how to carry out the task. Implicit memory is distinct from implicit learning in that it is characterized as the influence of the previously memorized piece of information on a task without the explicit or deliberate attempt to recall the memory (Underwood & Bright, 1996).

[2] The benzodiazepines enhance the action of the neurotransmitter, GABA (Gamma Amino Butyric Acid). Neurotransmitters are chemicals which enable the brain cells to transmit impulses from one to another. They are released from brain cells by electrical signals. Once released, the neurotransmitters signal inhibition or excitation of neighboring brain cells (for review, see Szara & Lutford, 1981).

2.4.2 The Neural System That Mediates Familiarity Memory

There is a current debate about whether recollection and familiarity depend on the same or on different psychological processes and, by implication, on the same or different neural mechanisms. Starting from the idea that the increasing levels of felt familiarity should modulate activity in brain structures mediating familiarity memory, Montaldi, Spencer, Roberts and Mayes (2006) tried to identify the neural system that underlies scene familiarity memory.

Recollection was found to activate the hippocampus, and left anterior and inferolateral frontal and parietal cortices more than strong familiarity. In contrast, no brain region that was unaffected by recollection was modulated by variations in familiarity strength.

The fMRI data suggest that familiarity memory for pictured scenes is mediated by a system of interlinked brain structures both within and outside the medial temporal lobe (MTL). It is claimed that perirhinal cortex[3] lesions are critical for familiarity but since selective lesions of this cortex rarely occur in humans, it is unknown whether they are also necessary for recollection (Montaldi, Spencer, Roberts & Mayes, 2006).

Buckner and Wheeler (2001) have proposed that familiarity feelings may be generated by the parietal and frontal cortex structures that show memory-success-dependent activation for many kinds of remembered information. These structures may process the changed inputs from rerepresenting structures to produce memory feelings for recollection as well as familiarity.

[3] The perirhinal cortex lies on the ventral surface of the temporal lobe in primates and in equivalent regions in all other mammals (Murray & Richmond, 2001).

In summary, a system of brain regions whose activation was sensitive to the perceived familiarity strength of studied scenes has been highlighted. The regions included the frontal, temporal, parietal, and retrosplenial cortices. Within the MTL, they included the perirhinal cortex, but not the hippocampus, which was only activated by relatively effortless recollection (Montaldi, Spencer, Roberts & Mayes, 2006).

2.5 Virtual Reality

A typical dictionary definition of the term *virtual reality* (VR) is "an image produced by a computer that surrounds that person looking at it and seems almost real" (Longman, 1995, p.1597). Here, the word *reality* refers to the *external physical world.* Reality is virtual, if the reality suggests something can be explored by our senses, and yet does not physically exist.

There is a general acceptance that "Virtual Reality is about creating acceptable substitutes for real objects or environments, and is not really about constructing imaginary worlds that are indistinguishable from the real world" (Vince, 1999, p.27). Although there are lots of other definitions for the technology, they all mean that VR is an interactive and immersive (with the feeling of presence[4]) experience in a simulated/autonomous world (See Figure 2-1).

---

[4] Presence is the subjective perception that a mediated experience seems very much like it is not mediated. (For further reading: Slater, M (2003). A Note on Presence Terminology. *Presence-Connect*, 3 (3).)

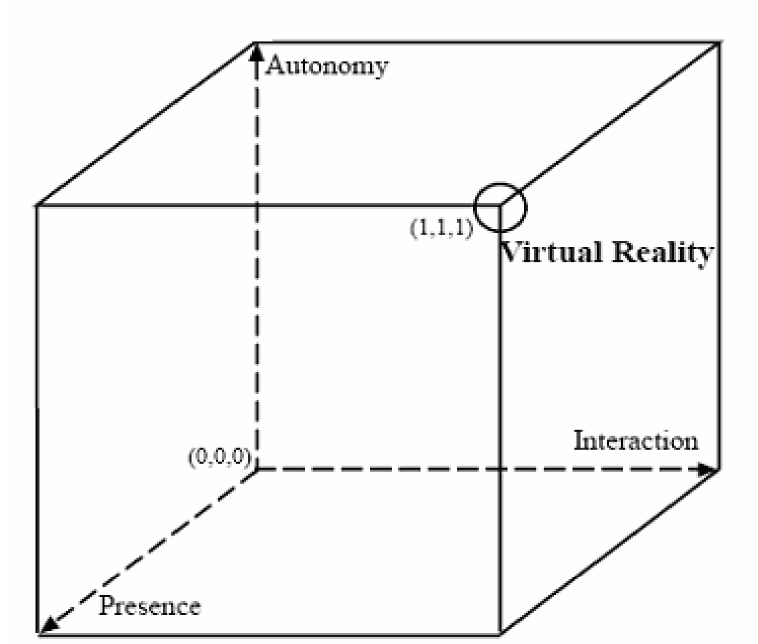


Figure 2-1 Autonomy, Interaction & Presence in VR (adapted from Mazuryk & Gervautz, 1996, p.4)

Virtual Reality Environments have three principal variants:

- Desktop Virtual Reality
- Augmented Virtual Reality
- Immersive Virtual Reality

When the 3D graphical virtual world is displayed on a standard computer screen, it is called Desktop Virtual Reality in which PCs and workstations can be used as screen-based Virtual Reality systems. This does not give true 3D depth perception and the sense of presence is low. The reason for this is that the user's peripheral vision is still in the real world, while using the standard PC. On the other hand, when the user has the sense of being immersed in the 3D virtual world by wearing head-mounted-displays and/or instrumented suits, the system is called Immersive Virtual Reality in which the user sees true stereo images and true 3D depth.

Between the "real" reality and "virtual" reality there some variations that include both types of realities. Ranging from the completely real through to the completely virtual, the Virtuality Continuum (See Figure 2.2) encompasses all possible variations and compositions of real and virtual objects that might be presented in a display.

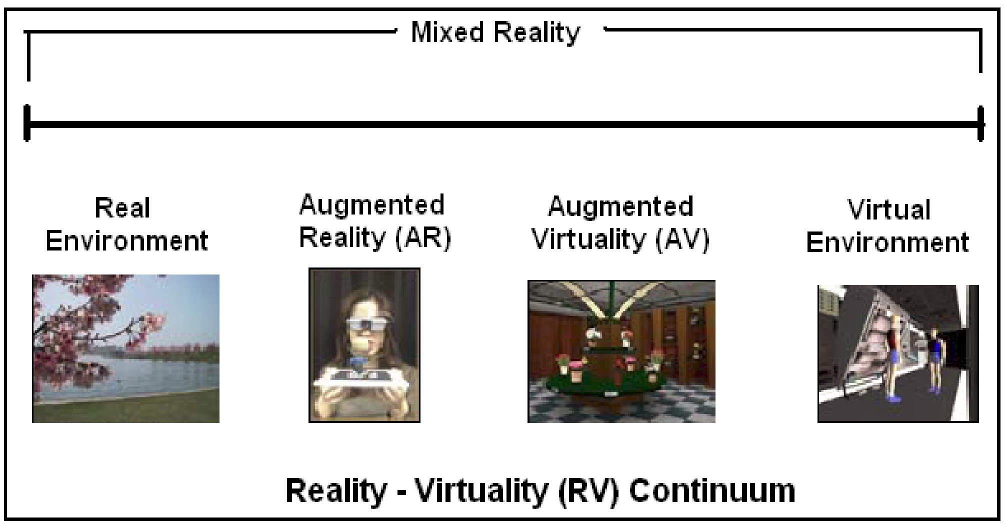


Figure 2-2 Milgram's Reality-Virtuality Continuum (Adapted from Milgram, Takemura, Utsumi & Kishino, 1994)

Both Augmented Reality (AR) and Augmented Virtuality (AV) combine real and virtual images in order to enhance the reality and provide the user with real-time interaction. AR adds virtual images into the real-world scenes. Most commonly display augmentation is achieved by using mirrors to superimpose computer generated graphics optically onto directly viewed real-world scenes (Milgram et al., 1994). On the other hand, AV adds "reality" to the video image. For example, a user's hand can be introduced into the otherwise principally graphic world, in order to point at, grab, or somehow otherwise manipulate something in the virtual scene (Takemura & Kishino, 1992).

The major variant of the Virtual Reality technology is the immersive one. The key elements in experiencing this type of virtual reality are (Sherman & Craig, 2003):

- Virtual world
- Immersion
- Sensory feedback
- Interactivity

Virtual world is an imaginary space where there is a collection of objects and some rules and relationships governing these objects. They are often manifested through a medium and they provide the user with images of 3D scenes for allowing him/her to navigate, explore and interact with them. In order to achieve this goal, real-time graphics are required because of the need for making the user believe that they are part of a virtual domain.

Immersion can be defined as the sensation of being in an environment. This term can be used in two ways: mental immersion and sensory immersion. Mental immersion is a state of being deeply engaged while physical one means bodily entering into a medium. Physical immersion is a defining characteristics of Virtual Reality; mental immersion is probably the goal of most media creators such as novel writers, film makers, animators, etc. (Sherman & Craig, 2003).

Unlike more traditional media, Virtual Reality allows participants to position their body and to affect events in the virtual world. Moreover, it allows us to purposefully reduce the danger of physical reality and to create scenairos not possible in the real world. For these purposes, “sensory feedback” is an ingredient essential to Virtual Reality (Sherman & Craig, 2003). This ingredient can be defined as the presentation of information about the virtual world to the participant’s senses. For accuracy, the system should provide direct sensory feedback to the participants based on their physical position. In this respect, two aspects have to be taken into account:

- the user’s location and motion in the real world
- the position of the user’s head and limbs

In Immersive Virtual Reality, users have the chance to move around and monitoring their absolute position is necessary for the reflection of their movement in the virtual world. Sensors, which are implemented by the technologies like infrared beams and ultrasonics, are used for tracking the user’s head position. This is important because of the need for correlating the user’s motions in virtual reality with the perceived change in the virtual world.

Virtual Reality Environments are highly interactive and, therefore, there is a need for many types of input and output technologies. The devices that are used in virtual environments are truly interactive because they combine multisensory feedback with input from the user. In general, a 3D mouse, instrumented gloves and suits are used as the input devices, which allow the user to navigate or to pick objects and communicate hand gestures to the host software within a Virtual Environment (VE). Furthermore, glasses and displays such as 3D screens, Head-Mounted-Displays, retinal displays, CAVEs (rooms that display the virtual environment), panoramic screens, virtual tables and augmented realities[5] are used as the output devices.

Many different cognitive factors play roles in the use of VE applications. These include issues related to perception, attention, learning and memory, problem solving and decision making, and motor cognition (Munro, Breaux, Patrey & Shelton, 2002). A number of characteristics of presently available virtual environments that raise some issues for perception are:

- very poor resolution of displays,
- problems with alignment and convergence,
- low-quality 3D models presented on displays.

These characteristics place substantive constraints on the role of bottom-up data processing while perceiving virtual environments, and as a result they increase the importance of the top-down contributions of expectations and experience (Munro et al., 2002).

Since virtual environments are useful for the study of spatial cognition, it has also been used as a tool for studying the abilities of participants for navigating and path finding in explored virtual environments (Darken, 1999; Bowman, Davis, Hodges & Badre, 1999; Stankiewicz et al., 2004; Arthur & Hancock, 2001; Witmer et al., 1996; Richardson, Montello & Hegarty, 1999; Moffat, Zonderman & Resnick, 2001; Waller, 2000; Üke, 2005). Individual differences also gained attention while studying Virtual Environments in order to design better systems for the users that have different cognitive abilities (Chen, Czerwinski & Macredie, 2000; Waller, 2000). The findings showed that the

[5] The real scenes that are enhanced or "augmented" with computer graphics.

psychometrically assessed spatial ability and proficiency with the navigational interface makes substantial contributions to individual differences in the ability to acquire spatial information from a virtual environment. The effect of gender was also examined and it was found to influence many virtual environment tasks, primarily through its relationship with interface proficiency and spatial ability (Waller, 2000).

Research and experimental applications of Virtual Environments (VE) have clarified some cognitive aspects of VE usage but many questions have been raised for further research. For example, a question that has not yet been resolved is the extent to which perceptual experience with VE representations is required to ensure accurate perceptions. Another important question is the extent that the special cognitive demands of the VE experience interfere with learning. The special advantages or disadvantages offered by VE-based learning for conveying different types of knowledge is also unknown.

Experimentation in psychology entails a tradeoff between experimental control and ecological validity. Virtual Displays afford less of a tradeoff than do traditional approaches to psychological experimentation (see Figure 2-3).

Furthermore, especially immersive virtual displays provide us with ecologically valid experiments, where the experimenter has the chance to maintain complete control of the virtual world around the subject (Loomis, Blascovich, Beall, 1999). Ecological validity helps us to generalize from observed behavior in the laboratory to natural behavior in the world (Schmuckler, 2001). In this study, this valuable attribute of virtual reality helped us to control and manipulate the experimental environments while observing perceptions of the participants in a more natural way than classical experiments in the literature.

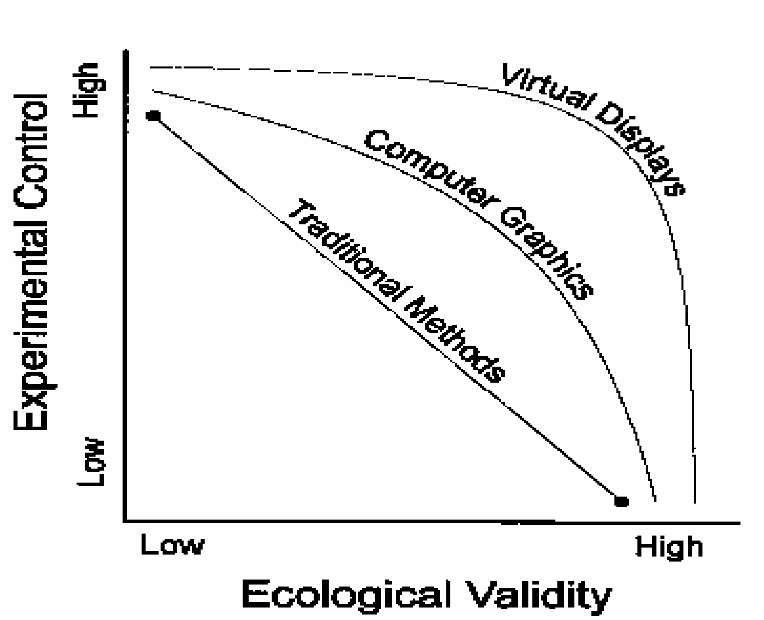


Figure 2-3 Tradeoff between experimental control and ecological validity (adapted from Loomis, Blascovich, Beall, 1999, p. 558)

Some aspects of VEs are also have potentially important consequences for the role of attention processes during VE experiences. We know about what the user is currently attending in these environments with the help of the eye movement observation systems attached to the VE equipments. This attribute of VEs was also used in this study in order to examine the gaze of the participants according to their cognitive needs during a process. Furthermore, this information was used to have insights about those consequences that results from the usage of 3rd dimension in the Virtual Environments.

2.6 Eye-Tracking

The visual world contains more information than can be perceived in a single glance. Consequently, one's perceptual representation of the environment is built up via the cognitive processing of this visual information. In an attempt to better understand this process in humans, the eye-tracking technology provides a useful methodology (Koesling et. al., 2001). When we scan a visual scene, such as a picture, our eyes alternate between rapid jumps and brief stops. These rapid jumps and brief stops are called saccades and fixations, respectively.

The saccadic eye movements are fast, ballistic movements of the eye. They are used to bring a new peripheral target into the fovea[6]. The term comes from an old French word meaning "flick of a sail" (Gregory, 1973). A person performs thousands of saccadic eye movements for many different activities every day. For example, while reading the gaze jumps from one word to the next one, each of these eye movements being a saccade (Casaña Pérez, 2004). Whereas saccades usually are made from one target to another in space, saccades can be executed with high degree of precision to special positions defined cognitively such as signals in other modalities, verbal commands, memories of spatial locations. For example, when a person is told to look at a certain object in a scene h/she can execute several saccades towards the location of that object which is defined with the verbal command. Whereas this suggests voluntary control, aspects of saccades (direction, velocity or amplitude) cannot be changed after they are initiated (May & Badcock, 2002). Two mechanisms are thought to govern saccadic eye movements: a spatial system, which tells the

[6] The fovea is the small region of highest acuity around the optical axis and the rest of the retina is called periphery (Dikici, 2004).

eye *where* to go and a temporal mechanism that tells the eye *when* to initiate a saccade (Velichkovsky, Dornhoefer, Pannasch & Unema, 2001).

When observers look at objects in the environment they position their eyes so that the objects falls on the area of the retina that has the best visual acuity[7]. When this is achieved the observer is said to have fixated the object. With a stationary target, fixation stability may be defined as the degree to which the eye is stable with reference to some fixation point on the object (May & Badcock, 2002).

Since the motion of images across the retina is fast, little information is processed during saccadic movement. On the other hand, eye fixations enable us to focus our attention like a spotlight. Therefore, one way of exploring what people pay attention to in any given situation is to use a computerized eye-tracking system to track the location and duration of their fixations (Moran, 2000). The analysis of visual fixations with these systems can provide rich information about the user's attention. It was shown that it is possible to distinguish between preattentive scanning and cognitive elaboration using parameters of fixation durations. For example, the eye movements and risk perception were analyzed during a driving task performed on a PC-based driving simulator. After the analysis, the fixational behaviour during critical events showed that a shift in the processing level from preattentive to attentive is recognizable on a very short, phasic time scale. A sudden increase in fixation duration was observed upon detection of a critical event (Velichkovsky et al., 2001).

[7] Visual acuity (VA) is acuteness or clearness of vision, especially form vision, which is dependent on the sharpness of the retinal focus within the eye, the sensitivity of the nervous elements, and the interpretative faculty of the brain (for further information see Cline, Hofstetter & Griffin, 1997 or Haber & Hershenson, 1973).

During free viewing of a scene, fixation durations appear highly variable, ranging from less than 100 milliseconds to several seconds, and such changes even take place within consecutive fixations. This is in noticeable contrast to saccadic reaction times, which are relatively stable, usually ranging from 150 to 250 ms (Velichkovsky, et al., 2001).

The video-based eye tracker is the most suitable type of tracker for 2-dimensional recording of these eye movements of a participant who is relatively free to move about in a region of space. It is currently used for communication and control in the VE (Wilder, Hung, Tremaine & Kaur, 2001). The tracker captures a video image of the eye illuminated by a distant, low-power, infra-red light source, and creates an image that is seen as a highlight spot on the surface of the cornea. The image is analyzed by a computer, which calculates the centroid of the corneal reflection as well as the centroid of the pupil. The corneal reflection[8] from the front spherical surface of the eye is insensitive to eye rotations, but it is sensitive to translational movements of the eye or head. On the other hand, the pupil center is sensitive to both rotation and translation. Thus, the difference between the pupil center and the corneal reflex provides a signal proportional to eye rotation, and thereby the direction of gaze, that is relatively free of translational artifacts[9].

In head-mounted eye trackers, the infra-red light source, the camera that views the eye, and a second camera that views the observed scene, are all mounted on the head. Additionally, as the user roams about in a prescribed space, a magnetic sensor is affixed to the head to record the position of the user in the

[8] The light which is reflected on the surface of the cornea (See Yoo, Kim, Lee & Chung, 2002 for detailed information)

[9] Artifacts generated by small lateral (translational) movements of the eye or the head (Gur & Snodderly, 1997).

space. The system computes the point of gaze in the image obtained by the scene camera. After a simple calibration procedure, the tracker can be used to record eye gaze position in the monitor display (Wilder, Hung, Tremaine & Kaur, 2001). The calibration procedure in this study is explained in section 4.1.2.

In summary, the reason for using eye-tracking technology is that the data collected from the tracker can provide not only behavioral end products of such cognitive processes like visual change perception but also clues to the process through which they are achieved. Importantly, this sensitive measure can also be used in ways that do not interrupt task processing with requests for metacognitive reports or other overt responses. Therefore, eye-tracking allows for a certain degree of ecological validity in task performance, as the responses it collects are ones that typically occur regardless of experimenters' instructions and participants' intent (Richardson, Dale & Spivey, 2007).

CHAPTER 3

LITERATURE REVIEW

The following literature review provides a brief overview of the previous work on the subject of this study. For this goal, several topics are included, such as change detection and change blindness. Furthermore, eye tracking studies and their applications in Virtual Reality environments are also summarized in this chapter.

3.1 Change Detection Studies

The capacity limits of attention and working memory set fundamental constraints on the way that the information in visual scenes is processed. Numerous experiments reveal that only a small fraction of the information in an image can be attended during a brief presentation, and retained in working memory. There is general consensus that only the ”gist” of a scene can be apprehended in a brief presentation, along with 3 or 4 items, or “object files”, together with information about scene “gist”, and other higher level semantic information (Irwin & Andrews, 1996; Hollingworth & Henderson, 2002; Luck & Vogel, 1997). An important and relatively unexplored issue is how these limitations play out in the context of vision in the natural world. How do we ensure that attention is directed to the right place at the right time? Humans for the most part manage extremely well to see what they need to see, and avoid colliding with pedestrians or tripping over curbs. This seems quite remarkable given the limited bandwidth that attention sets on visual processing. This limit on how much visual information is remembered from a brief presentation suggests a similar constraint on how much visual information is remembered from eye fixation to fixation (Palmer, 1990).

It is often assumed that the solution to the problem of directing one's attention on the right place is that attention is attracted by salient stimuli or events in the visual array. There is evidence that high spatial frequency content, edge density, and local contrast play a role in attracting fixations (Mannan et al, 1997; Krieger et al, 2000; Parkhurst & Niebur, 2003; Reinagel et al, 1999). It has also been demonstrated that visual saliency, based on image features such as color, intensity, contrast, and edge orientation, can account for some of the variance in gaze distribution when viewing 2D images (Itti & Koch, 2000, 2001; Koch & Ullman, 1985; Torralba 2003). In addition, sudden onset stimuli have considerable ability to capture attention even if the observer's attention is directed elsewhere (Jonides & Yantis, 1988; Yantis, 1998; Theeuwes, 1994). Other stimuli, such as a unique color or shape, or motion stimuli may also capture attention (Yantis & Hillstrom, 1994; Franceroni & Simons, 2003; Chastain et al, 2002). Attentional capture may also be accompanied by oculomotor capture[10], where gaze is diverted to the novel event (Theeuwes et al, 1998; Irwin et al, 2000). There is some uncertainty whether novel stimuli invariably attract attention whatever the task (Yantis, 1998; Folk et al, 1992; Gibson & Jiang, 1998; Gibson & Kelsey, 1998;

[10] Subjects perform a search task while eye movements are monitored. Then, an additional distinctive item appears in the display. Here, capture is reflected by an inappropriate eye movement toward the irrelevant item (For detailed information see Theeuwes et al, 1998; Simons, 2000b).

Turatto & Galfano, 2000). However, Brockmole and Henderson (2005) found that abrupt onsets of novel objects attracted a fixation when participants viewed photographic images of natural scenes, regardless of whether they were explicitly told to search for a new object that might appear, or whether less specific memory instructions were given. These authors also found that new objects attracted gaze even when presented during a saccade, although less reliably and with longer latency than those presented during a fixation, as might be expected given the nature of the retinal transient signal produced when the eye is stationary.

Most of the work showing stimulus based effects of attentional or oculomotor capture have been done with 2D experimental displays and either simple geometric stimuli or photographic renderings of natural scenes. There are reasons to suppose that attentional and oculomotor capture in 2D environments might not be entirely effective in the context of natural visually guided behavior. Even when participants are instructed to inspect images of natural scenes, the situation does not accurately reflect the challenges of visually guided behavior in real, 3D environments. Acting within a scene is very different from inspecting 2D images. Instead of a single stationary image, a complex image sequence is generated on the retina as the observer moves through the scene. Sudden onsets are rare in real environments, and many kinds of information are important to the participant, in addition to sudden onsets. For example, observers need to be aware of irregularities in the pavement, or the location of obstacles. Motion can effectively capture attention when presented in an otherwise stationary display, but might be masked in natural environments where retinal transients are continuously generated by the observer's motion through the environment. Thus it remains problematical how observers distribute attention in an appropriate manner in natural vision.

Recent work in natural tasks has demonstrated that the observer's cognitive goals play a critical role in the distribution of gaze during ongoing natural behavior (Hayhoe & Ballard, 2005). In extended visuo-motor tasks such as driving, walking, sports, and making tea or sandwiches, fixations are tightly linked, step-by-step, to the performance of the task (Land & Lee, 1994; Land & Furneaux, 1997; Land, Mennie, & Rusted, 1999; Patla & Vickers, 1997; Pelz & Canosa, 2001; Land & Hayhoe, 2001; Turano, Geruschat & Baker, 2003; Hayhoe, Shrivastava, Mruczek & Pelz, 2003). Participants exhibit regular, often quite stereotyped fixation sequences that are tightly linked, in time, to the actions, and very few irrelevant areas are fixated (Land et al, 1999; Hayhoe et al, 2003). This does not entirely solve the problem of how gaze will be distributed during goal-directed natural behavior, however. In a stable environment such as a table top, a participant's behavioral goals can be achieved by fixating only the task-relevant objects. In other environments, however, such as driving or walking down the street, the goals are less well defined, and it is not always possible to anticipate what information is required. How does the visual system handle unexpected stimuli? There is some evidence that participants can handle unexpected events by monitoring the environment in a strategic manner determined by the context. For example, drivers monitor the neighborhood of intersections to locate stop signs, and walkers monitor other pedestrians' trajectories to avoid collisions (Shinoda et al, 2001; Jovancevic et al, 2006). Therefore, it can be suggested that observers can learn the dynamic properties of the environment in order to distribute gaze in an optimal manner.

A different kind of support can be suggested by the finding that long-term memory for scene content is quite good (Hollingworth & Henderson, 2002; Henderson & Hollingworth, 2004; Melcher & Kowler, 2001; Melcher, 2005).

Few scenes are entirely novel, and observers make large numbers of fixations within a given scene. For example, within 5 minutes an individual will have made on the order of 1,000 fixations. Over repeated exposures to real scenes, participants may be able to build up quite elaborate long-term memory representations. It is well known that observers are insensitive to changes in scenes made during saccades, or when masked by some other kind of transient such as a blank screen or a blink (Rensink, 2000; Simons, 2000a; O'Regan, 1999). This has typically been interpreted to mean that very little information is retained in short term visual memory from one fixation to the next as observers move their gaze around a scene (Henderson, 2001). However, the finding that there may be robust long-term memory representations has important implications for how changes might be detected. A detailed long-term memory representation of a scene may be useful in guiding participants' attention to changes. If the scene is efficiently coded in long-term memory, different mechanisms might be available for coding new information. Participants may compare the current image with the stored representation in a manner similar to that suggested by Rao and Ballard (1998), and a mis-match, or "residual" signal may serve as a basis for detecting a change. Such mis-matches might serve as a basis for attracting attention to changed regions of scenes. This may allow participants to be particularly sensitive when there are deviations from familiar scenes, and thus attention may be drawn to regions that do not match the stored representation. Evidence for this was found by Brockmole et al. (2005). When participants were given 15 seconds pre-exposures to images of natural scenes, new objects were able to attract gaze in a subsequent brief exposure, even when the object was presented during a saccade, and there was no retinal transient associated with its appearance, and no opportunity to construct a short term memory representation. The authors suggest that the pre-exposure allowed participants to construct a long-term memory representation of the scene, as a basis for discriminating the new object. Thus when the scene is familiar, changes may be more readily detectable.

Another evidence for the importance of familiarity with the scene comes from work by Droll et al (2004). These authors made changes to the color of one of a set of red and blue virtual blocks that the participant was required to sort. When a block changed color from red to blue, or vice versa, participants were relatively insensitive to the color change. Changing the block color to one not present in the scene, such as green, was immediately apparent. This was true whether or not other blocks were present in the current view. This supports the attention-getting power of a stimulus that differed from the scene that observers were familiar with in the experimental context, which contained only red and blue blocks.

3.2 Change Blindness Studies

The change blindness phenomenon has generated a great deal of interest in the psychological, philosophical, and vision science literatures, both because it is strikingly counterintuitive, and because it undermines the traditional view that the visual system constructs a complete representation of the external world (Hollingworth & Henderson, 2004). Although failures to detect changes have been studied for decades (Rensink, 2002), recent work has brought change blindness into the realm of a typical perceptual experience (Simons & Rensink, 2005). The shift to a more general view of change blindness was triggered by the discovery in 1991 that observers often failed to notice large changes to photographs that were made during a saccade (Dennett, 1991; Grimes, 1996).

The experiments on “change blindness” can be grouped under three main categories (Noe, Pessoa & Thompson, 2000):

- the ability of perceivers to detect changes in photographs of natural scenes is greatly impaired when the changes occur during saccades (Grimes, 1996).

- the ability of perceivers to detect changes can be prevented by disrupting the visual system’s ability to respond to the motion transients produced by the changes (Rensink, O’Regan & Clark, 2003).

- change blindness is not related to the passive viewing conditions typically employed in the laboratory setting (Hayhoe, Bensinger & Ballard, 1998; Simons & Levin, 1997).

Grimes (1996) conducted some saccadic experiments to study integration accross eye movements. Participants were instructed to look for changes in images and push a button if they detected any change. The changes were introduced only during the transit of a saccade. A sample of the types of image changes presented to the participants as they viewed the scenes and their detection failure percentages can be listed as:

- A prominent building in a city sky becomes 25% larger. (100%)
- Two man exchange hats that are of different colors and styles. (100%)
- In a crowd of 30 puffins, 33% of them are removed. (92%)
- In a marketplace, brightly coloured fruits switch places among their four respective baskets. (75%)

- A parrot, comprising roughly 25% of the picture space, changes from a brilliant green to an equally brilliant red. (18%)

Under the light of these results, Grimes (1996) assumes that there is a point of contact between the previous mental experiences of the person and the information gathered from the current environment. In other words, the content of the image is extensively linked to a semantic description and participants are insensitive to changes that affect that description in the subsequent fixation while not changing the meaning of the scene.

Rensink, O'Regan and Clark (2003) conducted some experiments using the 'flicker' task, in which an original and modified scene alternate repeatedly, separated by a brief blank display, until observers find the change. Observers eventually found most changes, but it took an unexpectedly long time to do so, even when changes were large and made repeatedly. During the experiments, changes are easily identified when a valid verbal cue is provided, showing that poor visibility is not the cause of this difficulty (Rensink, O'Regan & Clark 2003). Changes are also easily identified when made to objects mentioned in brief verbal descriptions of the scene. Taken together, the authors suggest that we do not build up a representation of a scene that allows us to automatically perceive change even when sufficient viewing time has been given. The results also show that the changes to semantically central items are detected faster than changes elsewhere, suggesting that objects in a scene that preferentially receive attention are more likely to be encoded and compared. Therefore, the final claim of the authors is that "attention" is needed for change perception.

Rensink et al.'s (2003) flicker paradigm was also used to demonstrate attenuated change blindness for exogenously attended[11] items. The paradigm applied to scenes which were composed of arrays of simple line drawings, where one item on every trial was either replaced with another item, or was flipped about both its horizontal and vertical axes (Scholl, 2000). One item on each trial also exogenously captured attention via a colour singleton or a late onset, but this item was no more likely to be the changed item than any other, and observers knew this. When discussing the results of the study, the author suggested that the manipulations captured attention in the flicker paradigm and those changes to late-onset items and colour singletons were noticed faster because they were being attended. Furthermore, another study claimed that attentional effects exist and they are even stronger when the changes are unexpected (Simons & Levin, 1997).

The influence of semantic consistency was also examined by Hollingworth and Henderson (2000) and their results showed that a change to an object is more easily detected when that object is semantically inconsistent with its scene than when it is semantically consistent. The authors claimed that semantic properties of a scene region influence whether the representation of that region is or is not retained across views of the scene, and thus that the internal representation generated from a complex scene is not a veridical copy of that scene.

Hollingworth, Schrock and Henderson (2001) also used the flicker paradigm to examine the role of fixation position within the scene. In the modified image, a

---

[11] Attention which is guided by reaction to external stimuli (See Berger, Henik & Rafal, 2005 for detailed information).

single target object was changed either by deleting that object from the scene or by rotating that object 90º in depth (See Figure 3-1).

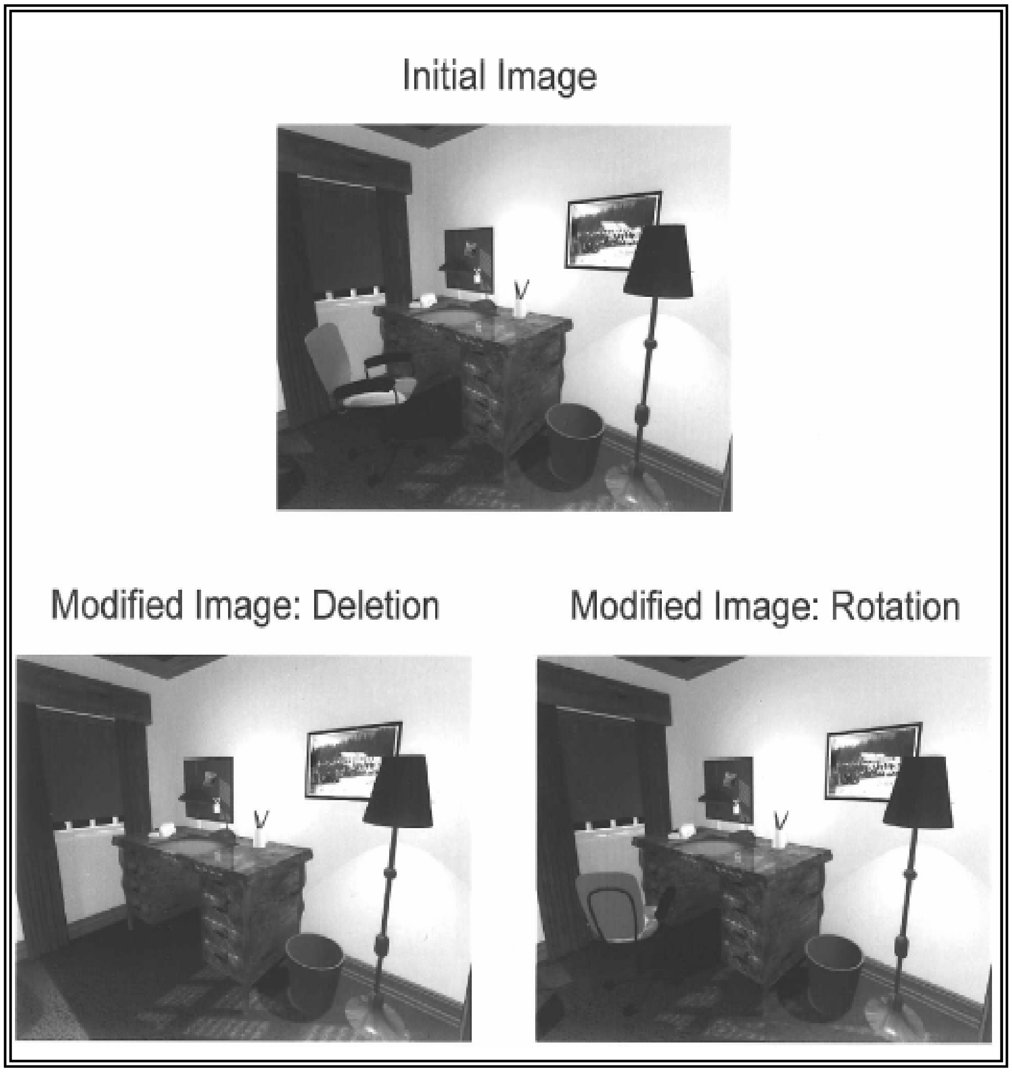


Figure 3-1 An example of a stimulus scene and the change conditions. The desk chair is the target object. This object was deleted from the scene between the initial and modified images in the deletion–addition condition. This object was rotated 90º around its vertical axis between the initial and modified images in the rotation condition (Hollingworth, Schrock & Henderson, 2001).

Their results indicated a significant causal role for fixation position in the maintenance of information across discrete views of a scene, leading to the detection of changes. Fixating the changing object appeared particularly important for the detection of more difficult changes in the rotation condition. In addition, these data have implications for the hypothesis that change detection in the flicker paradigm depends on the allocation of visual attention to the changing region (Rensink, 2000; Rensink et al., 2003; Wolfe, 1999).

Fernandez-Duque and Thornton (2000) also used modified change blindness tasks in their experiments in order to demonstrate that sensitivity to change does occur in the absence of awareness. In four experiments, they examined whether changes which are not explicitly detected by an observer can nonetheless influence subsequent behavior. Their findings suggested that focused attention can serve not only to *originate* representations of change but also to *modulate* representations that originate elsewhere in the visual system.

Another series of experiments were conducted to determine whether salience measurements applied to regions of pictures of outdoor scenes could predict the detection of changes in those regions (Wright, 2005). The detection and localization of changes was measured in a series of 40 picture pairs of natural scenes. Predictions about the detectability of changes from low-level image properties were found to be unsuccessful, but the rated salience of image points was claimed to be a good predictor. Object changes were found to be both more salient and more detectable than changes in shadows or colors. The author claimed that his results were consistent with the idea of a single topographic saliency map with a strong input from high level scene properties.

There is also a term "change blindness blindness" which implies the participants' unawareness of their inability to detect changes during the visual process. Beck, Levin and Angelone (2007) explored whether this error is related to participants' beliefs about the roles of intention and scene complexity in detecting changes. Their results showed that adults do not fully understand the role of intention and scene complexity in change detection. Therefore, the authors suggested that this metacognitive failure is most likely occur when the visual environment is complex and the viewer is not actively trying to detect the disappearance of objects.

Finally, if we think about the contributions of change blindness literature on change detection, we can claim that change detection is enhanced when attention is exogenously cued to the change region (Scholl, 2000) or when attention is directed to the changed region due to its perceived salience (Wright, 2005) or semantic importance (Rensink et al, 2003).

3.3 Using Eye-Tracking in Psychological Studies

Interests in visual perception resulted in the development of a new research environment, eye-tracking. In this research environment, researchers typically analyze eye movements in terms of fixations and saccades. Common analysis metrics include fixation or gaze durations, saccadic velocities, saccadic amplitudes, and various transition-based parameters between fixations and regions of interest (See section 1.1.4 for detailed information).

In the first direct study of eye movement patterns during scene perception, Buswell (1935) recorded the eye movements of viewers while they examined pictures of artwork, including pictures of complex buildings and sculpture (as cited by Henderson and Hollingworth (1999)). Buswell asked a number of questions: What are the regions that attract attention, what is looked at first, what is looked at last and what is looked at longest. In other words he investigated how attention is deployed. The results revealed that eye movement patterns were highly regular and related to the information in the pictures. These data thus provided some of the earliest evidence that eye movement patterns during complex scene perception are related to perceptual and cognitive processing.

The "eye movement-contingent display system" developed by McConkie and Rayner, exploited the phenomenon of saccadic suppression[12] as a tool for the investigation of various perceptual and cognitive aspects of reading (McConkie & Rayner, 1975). This eye-tracking system links an eye-tracker to a computer controlling a text display. This allows the experimenters to change the scene during saccades. As the eyes begin to move, their target destination is predicted and the text in that location is changed before the eyes arrive. This experimental setting allows the study of the influence of a variety of factors upon reading performance without directly disrupting the reading process (Grimes, 1996). In another reading study, eye tracking technology was used for the development of a general-purpose gaze-assisted translation aid for the texts written in a foreign language (Hyrskykari, Majaranta, Aaltonen & Raiha, 2000). They made use of information obtained from reading research, a language model, and the user profile in order for the system to accomplish the task.

[12] A decrease in perceptive ability associated with saccadic eye movements. For more information see Yu & Lee, 2000.

Velichkovsky et al. (2001) used eye-tracking technology to examine and compare quantitative and qualitative changes in fixation parameters due to task conditions, in particular with respect to preattentive search vs. attentive elaboration of critical events during driving. They conducted a study on eye movements and risk perception during a driving task that was performed on a PC-based driving simulator. The results showed that analyzing fixation durations has implications both for theoretical and applied research. The distinction between preattentive scanning and attentive elaboration in visual search seems to be supported by the data.

Top-down control of visual attention in real world scenes was also examined by monitoring the eye movements of human observers while instructed to search for a specific object in real-world scenes (Oliva, Torralba, Castelhano & Henderson, 2003). By using the results of this experiment, the authors generated a computational model which uses the statistical correlations that exist between global scene structures and object properties to define a region of interest in the image that is relevant for solving a task. The results showed that the model and human observers used the same region of interest in more than 85% of the cases. Therefore, they claimed that top-down information from visual context modulates early saliency of image regions during the task of object detection.

Cultural effects on eye movements were also investigated by measuring the eye movements of American and Chinese participants while they viewed photographs with a focal object on a complex background (Chua, Boland & Nisbett, 2005). Their findings demonstrated that eye movements can differ as a function of culture. In particular, the authors claimed that the East Asians attend to context while Westerners attend to objects instead of the general context.

Recent eye movement research has focused on extended visuo-motor tasks such as sports, driving, walking, building model rockets, washing hands, and making tea or sandwiches (Land & Lee, 1994; Land, Mennie & Rusted, 1999; Land & McLeod, 2000; Peltz, Canosa & Babcock, 2000; Shinoda et al, 2001; Hayhoe et al, 2003). These studies have found that the eyes are positioned at a point that is not the most salient, but is relevant for the immediate task demands. Fixations are tightly linked in time to the evolution of the task, and very few fixations are made to regions of low interest regardless of their saliency (Hayhoe et al 2003; Land et al 1999).

3.3.1 Eye-Tracking Studies in VR

In virtual environments, tracking allows proper rendering of images from the user's point of view. For example, head tracking provide an advantage by maintaining motion parallax cue[13], which improves the depth perception. Furthermore, one more important aspect can be taken into account: the visual acuity of the eye changes with the arc distance from the line-of-sight (Mazuryk & Gervautz, 1996). The primary goal of early eye trackers was to support research in human visual data acquisition. But as instrumentation technology continued to evolve, it eventually led to applications in a variety of settings

[13] When an observer is in motion the visual scene surrounding the person is represented as a drifting image on the retinas of the observer's eyes. The drift speed on the retina depends on the relative distance of a given object in the image. If the object is close to the observer, the drift speed of this object on the retina will be faster than when the object is further away from the observer. This relative motion of the visual image on the retinas is known as motion parallax and it is used by the visual system to generate a sensation of depth (Faubert, 2001).

where understanding the human perception, attention, search, tracking and decision making are of great importance.

Several advances in technology such as new mobile eye trackers that can be used in natural environments, and the development of complex virtual environments, now allow investigation of active gaze control in natural tasks in controlled conditions (Triesch, Ballard, Hayhoe & Sullivan, 2003; Shinoda, Hayhoe & Shrivastava, 2001; Jovancevic, Hayhoe & Sullivan, 2004; Jovancevic, Sullivan & Hayhoe, 2005).

Duckowski et al. (2000) developed a binocular eye-tracking Virtual Reality system for aircraft inspection training. They tracked user gaze directions in real time and calculated gaze/polygon intersections in order to enable the comparison of fixated points with stored locations of artificially generated defects located in the environment interior. Recorded gaze locations were used for the comparison of the performance of experts to novices, thereby measuring the effects of training.

The behavior of human observers performing visual search of natural scenes using gaze-contingent variable resolution displays were also examined in a desktop VR system (Parkust, Culurciello & Niebur, 2000). A two-region display was used where a high resolution region was centered on the instantaneous center of gaze, and the surrounding region was presented in lower resolution (See Figure 3-2).

Their primary finding was that reaction time and accuracy co-vary as a function of the central region size. A secondary finding was that fixation duration varies as a function of central region size. They reported that, participants tend to spent more time examining each fixation for small central region sizes than under normal viewing conditions. For large central regions, fixation durations were reported to be closer to normal. Therefore, the authors concluded that variable resolution displays can save computational resources without significant behavioral consequences.

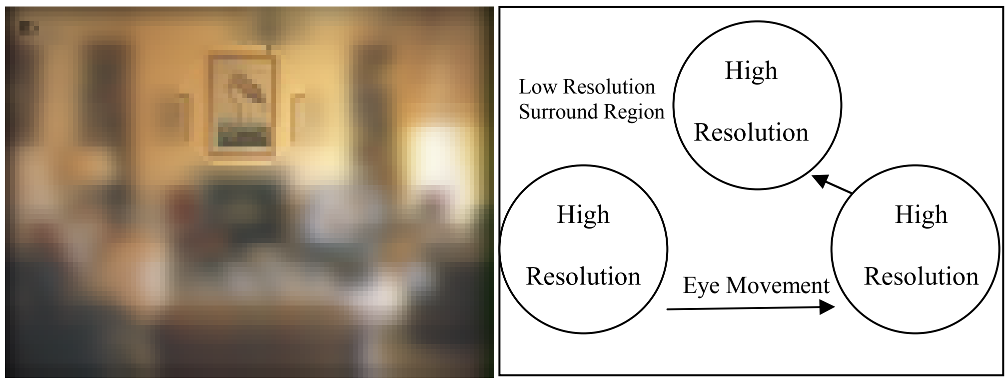


Figure 3-2 Variable Resolution Display (Parkust, Culurciello & Niebur, 2000, p.105)

Shinoda et al. (2001) conducted some experiments with virtual driving. In these experiments, participants' ability to detect "Stop" signs was examined. These signs were only visible for restricted periods of time and the appearance location of the "Stop" signs were either in the intersection or in the mid-block. Results of the study showed that participants' performance in detection of "Stop" signs was heavily modulated both by the instructions and the local visual context and the signs were invariably detected when at an intersection, even for a brief period. Furthermore, participants spend nearly half of the time deploying gaze in the region of the intersection, when required to drive normally. Therefore, the authors concluded that fixations on "Stop" signs were primarily controlled on the basis of active search according to an internally generated schedule, and this schedule depends both on the observer's goals and on learnt probabilities about the environment.

In a more recent study, Jovancevic et al. (2004) tested the hypothesis that attentional control is determined primarily by top-down factors, and that the distribution of attention is sensitive to the expectations the observer has about the current environmental context. Participants walked along a footpath in a virtual environment. Fixation patterns were examined to measure sensitivity to virtual pedestrians on a collision path. With a probability of approximately 10 percent, a pedestrian on a non-colliding path changed onto a collision course for about 1 sec, after which it returned to its original, non-colliding path. The results showed that participants often failed to fixate pedestrians on a collision path, suggesting that potential collisions do not automatically attract a fixation, unless the observer is actively monitoring for pedestrians in peripheral vision. In addition, the authors claimed that this is consistent with models where attentional control is determined by the task, with a minimal role for bottom up factors.

In a succeeding study Jovancevic et al. (2005) increased the saliency of colliders by increasing their speed by ~25% during the collision period, compared to the previous condition where they only changed direction. About 10% of pedestrians on a non-colliding path changed onto a collision course for 1 sec, and then returned to the original, non-colliding path. They claimed if peripheral vision is constantly monitored, or the deviation attracts attention, this should be revealed by a fixation. In one condition participants were instructed to follow a virtual pedestrian leader, and in another condition, to walk at their natural pace. Results showed that participants are most likely to fixate pedestrians in the first 2 seconds after they appear, during which time, in non-leader trials, participants fixate normal (non-colliding) pedestrians about 58% of the time; speeding colliders are fixated about 88%, and non-speeding ones about 68% of the time. However, this difference was not maintained for the remaining 3 seconds that pedestrians are typically in the field of view. In other words, speeding colliders added power to attract attention only when observers were likely to fixate pedestrians in general. Thus, the authors concluded the way participants distribute their attention across a scene is determined by a relatively small number of behavioral goals with varying priorities.

3.4 Summary & Contributions of the Study

The limited bandwidth of attention places tight constraints on visual processing of natural, complex scenes. Image properties, such as contrast, edges and chromatic saliency can account for some fixations when viewing images of scenes (Itti & Koch, 2001; Parkhurst & Neibur, 2003; Mannan et al, 1997). Cognitive goals also account for many of the fixations in natural behavior (Land et al, 1999; Hayhoe et al, 2003; Hayhoe & Ballard, 2005).

These properties are applicable for normal, natural scenes. But, what if an unexpected change occurs in the environment? Is attention automatically attracted to a change in the scene? There are two types of answers for this question. First, some evidence indicates participants are very poor at detecting changes in scenes if the change is masked by a transient – a phenomenon called "change blindness" (Rensink, O'Regan, & Clark, 2003; Simons, Levin, 1997; Triesch, Sullivan, Hayhoe & Ballard, 2002). On the other hand, other evidence suggests that novel objects attract attention (Yantis, 1993; Yantis & Jonides, 1996).

Depending on the goals of the observer, several kinds of stimuli including:

- The abrupt appearance of a new object (Yantis & Janides, 1984)
- The sudden disappearance of an existing object (Theeuwes, 1991)
- Objects characterized by unique color & shape (Theeuwes, 1994)
- Movement (Franconeri & Simons, 2003)

have all been shown to capture attention. On the other hand, it is uncertain whether the same mechanisms that drive attention to new or unique objects in simple stimulus arrays also operate under more naturalistic viewing conditions (Brockmole & Henderson, 2006).

Hollingworth and Henderson (2004) investigated two memory problems in scene perception. The first is the short-term retention and subsequent integration of scene information across saccadic eye movements. The second is the accumulation of scene information over longer periods of time during the visual exploration of a natural scene. This latter work focuses on the nature of the information retained from previously attended objects and on the role of long-term memory in scene perception. He and his colleagues found that despite evidence of change blindness, detailed visual information is reliably retained in memory from previously attended objects. Robust implicit effects of change indicate that explicit change detection does not provide an accurate measure of the detail of visual scene representation. In other words, the authors argued that elaborate representations of scenes are built up in long-term memory. If so, participants may compare the current image with the learnt representation in order to detect a change and such representations might serve as a basis for attracting attention to changed regions of scenes. On the other hand, Wang and Brockmole (2003) stated that observers keep track of objects or places they are approaching (i.e. those they can see) and lose track of objects or places that they have passed (i.e. those they cannot see).

Furthermore, Brockmole & Henderson (2005a) examined whether long-term memory (LTM) can direct memory-guided prioritization of new objects in real world scenes. Their stimuli consisted of full-color photographs depicting 30 real-world scenes. Two photographs of each scene were taken, differing only in the presence or absence of a single object in the scene (see Figure 3-3 for examples).

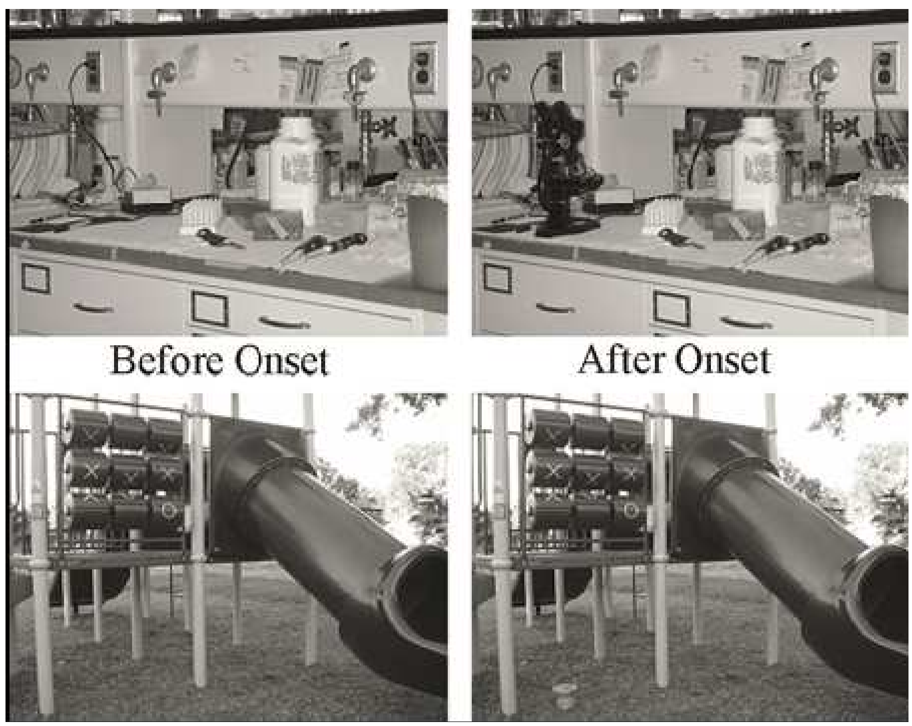


Figure 3-3 An example scene used in the study depicted both before and after the onset; in these cases the microscope and bucket (Brockmole & Henderson, 2005a, p. 859).

The results showed that observers can rely on their LTM to guide their attention through the scene & localize changes, even when sufficient time is not afforded to generate a short-term memory (STM) representation capable of guiding attention to the new object (Brockmole & Henderson, 2005a). Thus, participants should be more sensitive to changes in familiar environments than in unfamiliar ones.

In the present study, this issue was further investigated. In particular, it was asked whether the effect of scene familiarity generalized to natural environments. Does familiarity with scene content improve detection of scene changes in an immersive, 3D environment? As described above, it is necessary to observe active, visually guided behavior in natural environments because the stimulus conditions and cognitive goals are so different from viewing 2D images, even when those images represent natural scenes (Hayhoe & Ballard, 2005; Droll et al, 2005; Triesch et al, 2003). Since real natural environments are hard to control, an immersive virtual environment, where participants walked along a footpath in the presence of a variety of stationary objects, was devised. I examined whether the opportunity to become familiar with the environment influenced the distribution of gaze, and in particular, whether participants preferentially fixate scene changes when they have become familiar with the environment.

In order to understand the way that familiarity might improve detection of changes, I also manipulated the kind of changes that were made. Previous work has revealed attentional capture may occur with objects that vanish, as well as those that appear (eg Theeuwes, 1991), although offsets (i.e. disappearance of an existing object) might be less effective than onsets (i.e. appearance of a new object) (eg. Boot, Kramer & Peterson, 2005; Brockmole & Henderson, 2005). On the other hand, Mondy and Coltheart (2000) claimed that identification of object deletion was more likely than the identification of object addition. In addition to this debate about new objects, and objects that were removed, I also examined two more change types in order to be able to compare the nature of different changes in a broader context. The additional changes were movement of objects, and replacement of existing objects with different ones. This type of comparison was not made in the literature before and I think it will provide an important insight since both these factors are relevant in change detection manipulations, and may reveal the nature of the detection process.

Finally, I was interested in changes that did not engender a retinal transient. Since changes accompanied by a transient are typically easier to detect than those that are not (see, eg, Brockmole & Henderson, 2005a), this represents the most challenging situation for the observer, and, as described above, many of the situations where an observer needs to attend will not necessarily be accompanied by a transient signal.

CHAPTER 4

EXPERIMENTAL INVESTIGATION OF CHANGE PERCEPTION

Two different experiments were designed for understanding the general issue of what controls attention in virtually modeled natural scenes. In particular, it was investigated whether familiarity with the environment makes people more sensitive to changes in the scene. Furthermore, the effects of presence on change perception were also investigated by applying the same research design both in a desktop and an immersive virtual environment. For these issues, two different virtual environments were created. One of these environments was created by using immersive VR technology and 3D objects were explored by a Head Mounted Display (HMD) while the second environment was created with desktop VR technology and this time 3D objects were shown to participants as they explored the environment through an LCD monitor. In addition, the eye-tracking technology was used for each of the experiments in order to follow the position of the observers' eye-gaze, which helps to measure the attended location in the environment. In this chapter the research design of these two experiments and their results are discussed.

4.1 Experiment 1 – Immersive VR

The experimental design of this study includes participants walking along a footpath in an immersive virtual environment including both stable and changing objects while avoiding potential collusions with virtual pedestrians represented by simple colored 'robot-like' figures (See Figure 4-1). There were six pedestrians; two walking in the same direction and four in the opposite direction, at different speeds, so that their configuration varied continuously as the observer walked around the central monument.

4.1.1 Virtual Environment

In the environment, participants walked in a rectangular path around a monument (see Figure 4-3A). As the environment, a segment of a 3D model of a town (Performer Town) created by SGI was used, so that participants could walk down a virtual footpath in the town. The footpath includes four corners that correspond to walking along four sides of the 4.8x6m experimental room, a distance of about 29.6m. The dimensions of the virtual world are geometrically matched to the real world so that there is no visuo-vestibular or visuo-motor conflict generated by movement through the scene. In other words, the visual information gathered from the virtual world did not conflict with the participants' physical position in real world. For example, when a participant walked for 2 meters beyond his/her first position s/he saw the place where was 2 meters beyond of his/her first position in the virtual world.

In the environment, at each side of the monument there were 2 objects, making a total of 8 (see Figure 4.3).

Figure 4-1 A Snapshot from the Environment

These objects were divided into two sets: changing and stable objects. Changing objects were a gazebo, a dog and a firehydrant. In addition to these changing objects, there were also 3 unchanging objects (i.e. a water fountain, a street lamp, a house) which were located at the 2 sides of the monument where the changes occured. On the other hand, a trashcan, a mailbox, a billboard and another house were stable objects that were located at the other 2 sides of the monument which were not subjected to change (See as changing/stable sides in Figure 4.2 below).

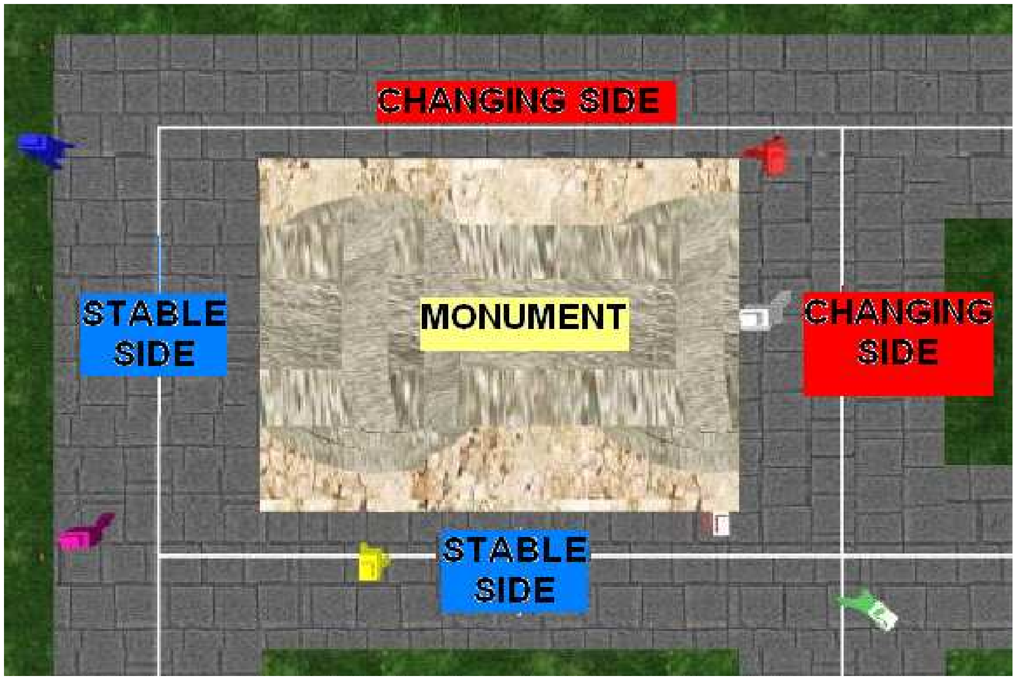


Figure 4-2 Bird's-eye view of the environment showing the changing and stable sides around the monument

For the first change, the gazebo was replaced with either a dog or a firehydrant. Here, two different objects, of approximately equal size, were used in order to test whether the results were influenced by the specific properties (i.e. color, shape) of the replaced object (see Figure 4.3B). The second change was the disappearance of this replaced object. This location remained unoccupied until the end of the trial (see Figures 4.3C & 4.3E). When a new object appeared it was placed in between two existing objects as shown in Figure 4.3D. That object then moved to a new location on the corner on the subsequent circuit around the monument (see Figure 4.3E). The new object that appeared in the environment was again either a dog or a fire hydrant. Half of the participants saw the dog as the new object (and the firehydrant as the replaced object) while the other saw the firehydrant instead (with the dog as the replaced object).

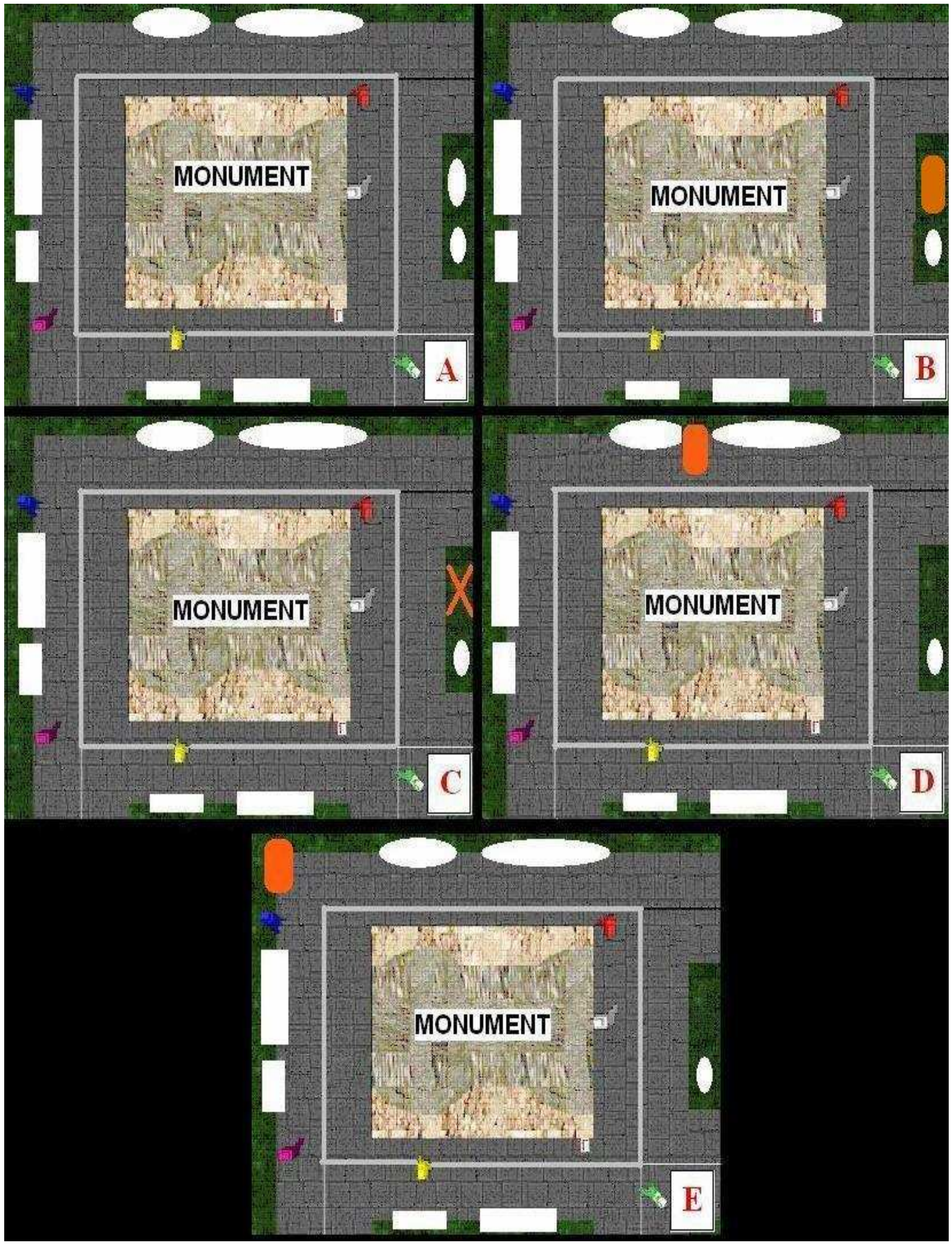


Figure 4-3 A. Bird's-eye view of the environment. White circles and white rectangles represent changing and stable objects respectively. B. An object was replaced with another object (all changes are represented with an orange rectangle) C. An object disappeared (the cross sign shows the previous location of the disappeared object) D. An object appeared E. The new object moved to a different location.

4.1.2 Physical Equipment

The experiments were conducted at the Virtual Reality Laboratory at the University of Rochester. This laboratory has several systems integrated to allow such a virtual reality experiment. For this study, I used a head mounted display with an eye tracker installed. In Figure 4.4, you can see the video based tracker for recording direction of gaze.

The eye position was calibrated by having the participant look at each of nine points on a 3 X 3 grid. The calibration was repeated between each trial to make sure that the noise of the eye tracker and the movement of the helmet on the head did not reduce the quality of the tracking. In addition, a video record of the scene, with eye position superimposed, was captured using a Hi-8 video recorder. An image of the left eye was superimposed on the video record at the top left corner to allow monitoring of potential track losses.

Figure 4-4 - V8 optics with ASL501 Video Based Eye Tracker (Left)

In order to allow the participants to walk a sufficient length, a Hi-Ball wide area motion tracking system is used to update the view inside the display while allowing the participant to walk approximately 30 meter perimeter of rectangular path in the lab (Figure 4.5). In order to remove jitter the Hi-Ball's position, information is filtered using an exponential equation:

$$Position(t) = 0.9*Position(t-1) + 0.1*New\_Data\ (t), \quad \text{(Equation 1)}$$

The HiBall tracker system is a recent development based on the Wide-Area Tracking research project of the Department of Computer Science of the University of North Carolina at Chapel Hill. It provides six-degree-of freedom tracking of devices in real time. An infrared-sensing subsystem is mechanically fixed to each device to be tracked. The HiBall view an environment containing a subsystem of fixed-location infrared beacons which is called the Ceiling. These subsystems are coordinated by a Ceiling-HiBall Interface Board (CIB) which provides communication and synchronization functions between the host computer and the attached subsystems (Welch et al., 1999).

Figure 4-5 First picture shows the Hi-Ball sensor and the second picture shows the Hi-Ball ceiling mounted positional array. Each unit contains 8 Infrared Emitting Diodes that are activated over a high speed network.

The system latency for updating the scene conditioned on a movement of the Hi-Ball motion tracking system was estimated as approximately 40-55 ms depending on when the data was received by the rendering computer. The update rate was sufficient such that participants usually did not experience a noticeable lag between head motion and the visual update.

Figure 4-6 Virtual Research V8 Head Mounted Display with 3rd Tech HiBall Wide Area motion tracker

The visual display was generated by a Silicon Graphics Onyx 2 computer at a rate of 60 Hz and was rendered in stereo on two LCD screens in the headset with 640 by 480 pixel resolution, and a visual angle of 48deg by 36deg.

4.1.3 Pilot Study

Before starting to conduct this experiment, some pilot tests were completed to form the final procedure. First, the possible number of turns in one trial was tested by having two lab assistants to walk in the experiment room. After completing this test and gathering the comments of the other experimenters that were using the same experimental room, the ideal number of turns for one trial was set to be 6. Here, the physical conditions of the users and the position of cables attached to the helmet and the trackers was taken into account. In other words, we tried to prevent participants from being motion or simulator sickness[14] while having the cables untangled. Then, the experiment was started to be conducted without having pedestrians in the virtual environment. After completing 5 sessions, a secondary task appeared to be needed since the dimensions of the virtual environment was small due to the physical constraints of the experimental room and, therefore, the possibility of fixating an object was high. In order to minimize the effects of this problem, pedestrians were added in the environment for making the walking process more challenging.

In the following sections you can have the information about the personal characteristics of the participants of this experiment and its final procedure which was formed after these pilot tests.

4.1.4 Participants

Participants were 38 students affiliated with the University of Rochester. 30 of them were undergraduate students while 8 participants were graduate students.

[14] Simulator sickness is the unwanted side effects which generally includes the symptoms of nausea, dizziness, headache or eyestrain and aftereffects which are the sense of balance, such as postural disequilibrium (Onay-Durdu & Çağıltay, 2006).

Experimental groups of this experiment were divided according to the familiarity level of the participants as *Experienced* for high-level familiarity and *Inexperienced* for low-level familiarity (See Section 4.1.5 for detailed description of the groups). In each experimental group there were 19 participants who were allocated to the groups randomly. Of the participants, 22 were female. The gender distribution of the groups can be seen in Table 4.1.

Table 4-1 Gender Distribution in the Groups

Gender / Group	Female	Male	Total
Experienced	11	8	19
Inexperienced	11	8	19
Total	22	16	38

The average age was 20,66 (sd = 2,822). The age range was from 18 to 35 (See Figure 4-7).

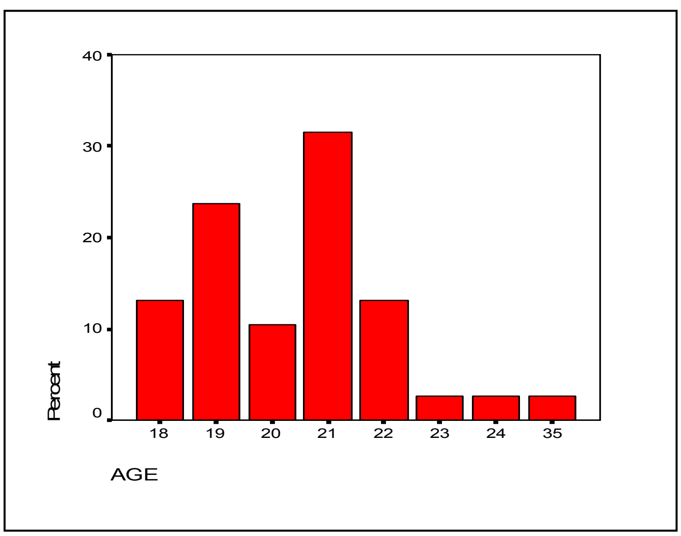


Figure 4-7 Age Distribution of the Participants

Besides the general questions of the study, I also searched for the possible effects of computer usage and game playing habits on participants' performance while analyzing the data. For this issue, participants filled a questionnaire including questions about their computer usage routine. Results showed that all of the participants were using computers at least for 5 years and most of them were using computers for more than 20 hours a week. Some of the participants used to play computer games (See Table 4.2). Moreover, all of the participants had normal or corrected-to-normal eyesight.

Table 4-2 The Distribution of Participants' Game Play Frequencies

Game Playing	Frequency	Percent
Never	22	57,9
Sometimes	15	39,5
Frequently	1	2,6
Total	38	100,0

4.1.5 Procedure

The experiment was conducted with 38 undergraduate and graduate students affiliated with the University of Rochester. They were paid 10 US Dollars for their participation. Each participant was given time to become familiar with the environment while walking the same path 6 times and this amount of walking was counted as 1 trial. In addition, the experiments were conducted in "one participant at a time" manner, so none of the participants was influenced by the other participants.

The experiment was done in the between-subjects format; two groups were defined according to participants' familiarity with the environment. Each group include 19 participants. The first group was labelled "*Inexperienced"* since they had no familiarization trials. On the other hand, the second group had 3 familiarization trials (i.e. 18 turns around the monument) in the experimental environment before the changes occurred and because of this reason they were labelled "*Experienced*". The participants were sequentially allocated to one of these two groups.

In this experimental setting, the distribution of gaze was examined during familiarization trials, and also following changes in the environment. As explained above, the 4 objects in the virtual environment (i.e. gazebo, water fountain, street lamp, house) were located at the 2 sides of the monument which are subjected to change. These changes that were applied to these objects include:

- appearance of a new object
- disappearance of an existing object
- movement of an existing object
- replacement of an existing object with a new object

The new objects that appeared in the environment were either a dog or a firehydrant. Here, two different objects were used in order to be sure that the results are not influenced by the property of the new object.

The *Experienced* group had a total of 4 trials: 3 familiarization trials and 1 trial when the changes occurred. The *Inxperienced* group had only 1 trial, when the same changes were made. The *Inxperienced* group walked around the monument once before any changes occurred. (Thus the *Inexperienced* group had a small amount of experience, resulting from the first circuit around the monument.) On the second circuit, the gazebo was replaced with either a dog or a firehydrant. On the next circuit, the replaced object disappeared. On the third circuit, a new object (a firehydrant if the replaced object was a dog, and vice versa) appeared in between the two existing objects. On the fifth circuit, this object moved to a new location that was previously unoccupied by an object. The same sequence of changes was followed for the *Experienced* group, but they occurred on the fourth trial, instead of the first.

There were 4 changes and they happened after the first round had been completed by the participants. Each change was visible for 2 rounds except for the disappearance of an existing object. This change could be perceived for 4 rounds since it happened during the second round and no other object appeared at the place of the disappeared object.

Before the experiment, each participant signed a consent form that explains the general rules such as they would be required to walk in the room, they could terminate the experiment whenever they want without losing the money they would get after the experiment, etc. Since they could not see the actual world during the experiments, the pathway in the laboratory was shown to convince the participants about the safety of the experiment. In order to be sure about the safety of the participants, a laboratory assistant walked with the participants during the experiments while holding the connection cables of the system (See Figure 4.6). The possible aftereffects of the usage of virtual environments were also explained and they were asked to inform the experimenter if they had any discomfort during the experiment (No participants reported discomfort).

After this information was given, they were located at the starting point of the environment while wearing the Head-Mounted Display (HMD). The equipment was adjusted according to each participant's physical condition (i.e. height, eye location, etc.) by using the nubs at the back and at the top of the HMD. The location of the eye-tracker was also adjusted to get a clear eye image on the screen. After these adjustments, the calibration of the eye-tracker was done as explained in section 4.1.2. Their task was explained as trying to become familiar with the environment while avoiding the pedestrians. This explanation was the same for both groups. The only difference between the two groups was the number of turns around the monument before the changes occured.

Total duration of one session (i.e. complete 4 trials) was varied between 15 and 25 minutes according to walking pace of the participant. After each session of the experiment completed, the participant was asked to report the changes that they had realized while walking in the virtual environment. These reports then used to understand the explicit knowledge of the participants about the changes that happened during the experiments.

4.1.6 Hypotheses

The results of this experiment test the following hypotheses:

H_1: Fixation durations for the *Experienced* group will be longer for the changes that occur in the experimental environment.

H_2: Fixation durations for the *Experienced* and *Inexperienced* groups will not be significantly different for the objects that do not change during the experiments.

4.1.7 Results

The logic of the analysis was calculating the average fixation time of a single object. The gaze durations were counted by analyzing the video results of the experiments frame by frame. Fixations were defined as a constant location within a 1 deg radius for a period of 100 msec or more. I was interested in the total amount of time fixating a particular object. This might be composed of a single long fixation or several shorter fixations on the object. I will refer to this as gaze duration to avoid confusion with the duration of a single fixation.

The location of eye fixation was taken as an index of the locus of attention within the scene. The total duration of all fixations on stable objects was found, and also on all changing objects, separated by category of change. Gaze was also divided up into fixations on the background and pedestrians. The total fixation duration was summed up for all stable objects, and also on all changing objects, and then these values were divided by the number of circuits around the monument, and the number of objects to give the total time fixating a single object in one circuit around the monument, averaged over objects. For the object that disappeared, fixation on the object's prior location was measured.

As explained above, there were two experimental groups namely *"Inexperienced"* and *"Experienced"*. First group had no familiarization trials. On the other hand, the second group had 3 familiarization trials in the experimental environment before the changes occurred. During the analysis, only the trial where the changes occured was analyzed (i.e. trial 4 for the *Experienced* group and Trial 1 for the *Inexperienced* group). All 6 laps in the trial were examined. The first lap was used only for calculating the average fixation durations for stable objects.

Although there were some stable objects on sides of the monuments where changes occurred, only those stable objects on sides where no changes occurred were included in the analysis, in case there was some influence of the nearby changes.

All analyses were done on the video images that were obtained during the experiments. These video images were then converted to a format that can be usable for PCs. After this convertion, the images were imported to the Windows Movie Maker Application and then obtained video records were examined frame by frame to find the fixations. Then, the duration of each fixation was calculated in milliseconds. For the analyses of the data from the first 10 experimental sessions, an expert on the field helped and explained the way the analyses should be done. After these analyses, the durations of fixations were summed up to get the numbers that are necessary for the calculations explained above.

Effects of Familiarity

After examining the data, average fixation duration for all changing objects for the *Inexperienced* group was found as 263 msec while that of the *Experienced* group was 534 msec. In addition, average fixation duration for all stable objects for the *Inexperienced* group was 330 msec and that of the *Experienced* group was 331 msec (See Figure 4-8).

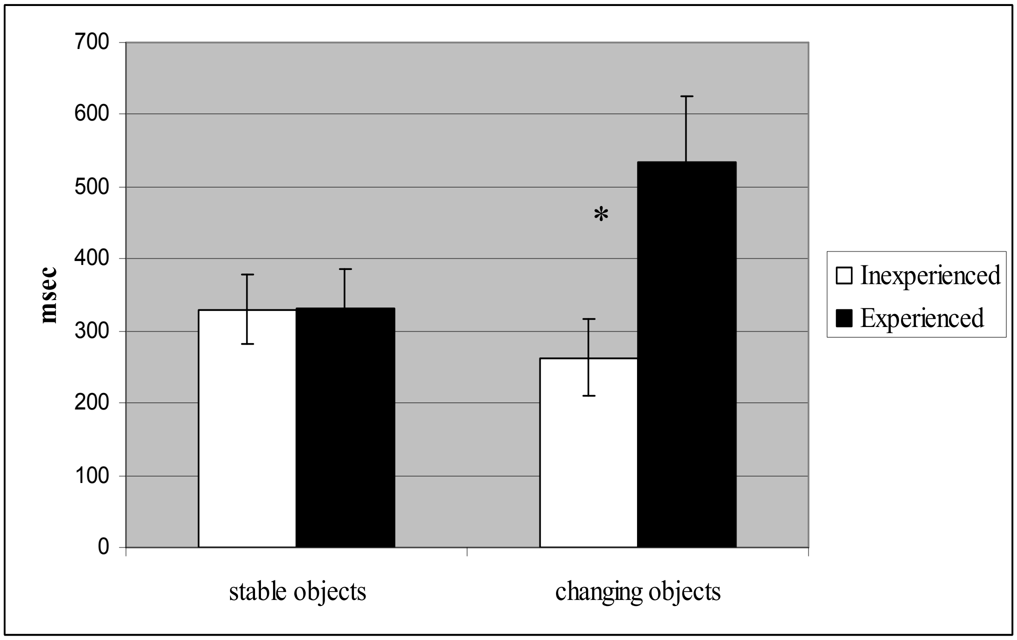


Figure 4-8 Total time gaze fell on an object on a given circuit around the monument, for stable and changing objects, for Experienced and Inexperienced groups. Error bars are +/- 1 Standard Error of the Mean across subjects

When the data was statistically examined according to both within and between-subject effects (2 (Experienced, Inexperienced) * 4 (stable, new, moved, replaced)), the results of mixed-ANOVA showed that the effect of group (*Experienced* vs. *Inexperienced*) was significant ($F(1, 33) = 4.219$, $p < .05$) for detecting these changes.

Since the fixation durations for the deleted object fell on an empty spot rather than an object as in the other change types, effects on this change type examined separately. The result of this analysis showed that there is a significant difference between the two groups according to their fixation durations for the deleted object ($F(1, 37) = 4.3$, $p < .05$).

Both groups had essentially identical fixation durations on the stable objects. However, participants who had experience in the environment looked substantially longer at the changing objects, an increase of approximately 170 msec. This suggests that prior experience in the environment indeed increased the likelihood that participants would look at a change in the scene. Note that the *Inexperienced* group's fixation duration on changing objects decreased slightly relative to the stable objects (matched pairs $t(18) = 1.84$, $p = 0.04$). This is probably a consequence of reduced fixations on objects that were removed or replaced (See below).

There were four types of changes that occured in the environment during the experiments. The effects of these different changes were also statistically examined. These data are shown in Figure 4-9 for the two groups. These changes and their representations (in italics) in the figure are listed below:

- appearance of a new object *(New)*
- movement of an existing object *(Moved)*
- replacement of an existing object with a new object *(Replaced)*
- disappearance of an existing object *(Deleted)*

When the within-subject effects of change type was examined, the results of mixed-ANOVA showed that the nature of change (new, moved, replaced) had no significant effect on fixation duration ($F(3,99)= .190$, $p = .903$). Here, again deleted object condition excluded in order not to influence the result with uncomparable (i.e. empty spot vs. object) data.

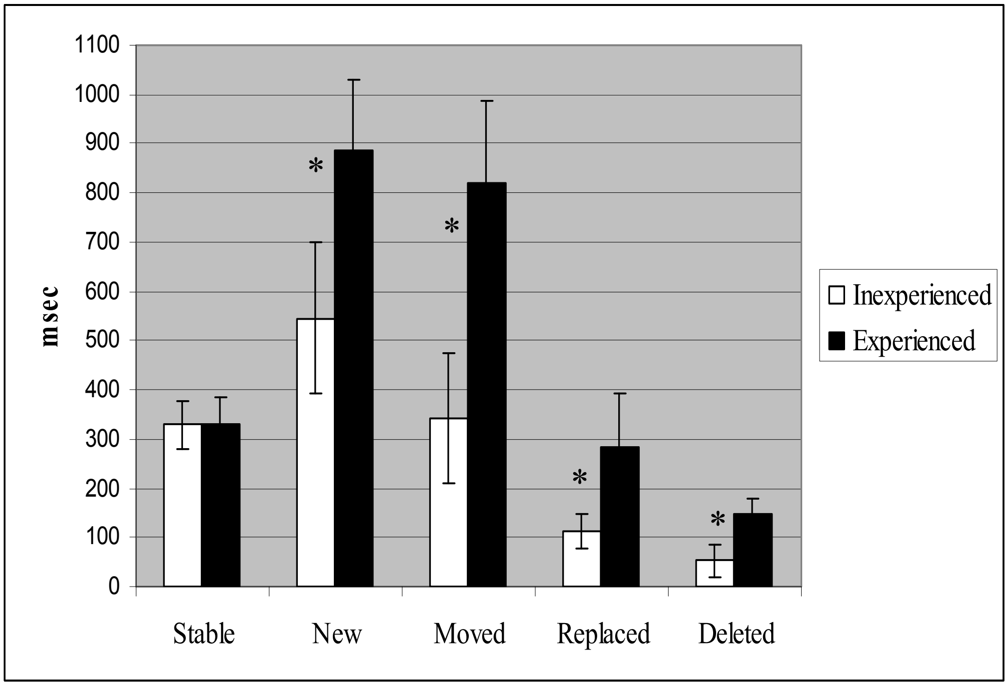


Figure 4-9 Gaze Durations for Different Changes

Experienced group fixated longer on every type of the changing object than the *Inexperienced* group. Numerically, the *new-object* condition was the most powerful change type to attract the gaze of the participants in both groups. In this manner, *moved, replaced* and *deleted-object* conditions followed the *new-object* condition in a descending way while the *deleted-object* condition was the least powerful change type.

The biggest difference between the average durations of fixations of the two groups was for the *moved-object* (478 msec). Furthermore, the differences between the average durations of fixations of the two groups on *new*, r*eplaced* and *disappeared* objects were 341msec, 169msec and 94msec, respectively (See Figure 4-9). Note that no differences were observed depending on whether the dog or the fire hydrant was used as a novel object, so the data were collapsed over these conditions.

Individual Differences

In each experimental group there were 11 female and 8 male participants in a range of ages between 18 and 35. Even though the study did not aim to explore the correctness of a hypothesis that includes individual differences, as an additional finding, gender and age differences were examined.

First, t-test was used to examine the effect of gender on participants' performance and no significant difference was found between males and females according to their fixation durations on changing objects ($t(36)= 0.423$, $p = .675$).

The correlations between the ages of the participants and their fixation durations on changing objects were also examined, and again no correlation was found to exist (Pearson's $r = -.041$; $N=38$; $p= .809$).

Finally, departments of the participants were categorized as natural sciences or social sciences and the correlations between these categories with the participants' fixation durations on changing objects were examined, and again no correlation was found to exist (Pearson's $r = .252$; $N=38$; $p= .127$).

Gaze Distribution

I also examined how participants distribute their gaze in the environment. The location of the fixations was classified into fixations on the path, surrounding environment (for example the grass, the monument, or in the distance), pedestrians, or objects, either changing or stable. The proportion of time spent fixating on each of these regions is plotted in Figures 4-10 & 4-11 for the *Inexperienced* and *Experienced* groups respectively.

Most of the fixations (54%) were located on the walking path for both Experienced and Inexperienced groups. This may reflect the ongoing demands of walking and staying on the path.The *Inexperienced* group devoted 30% of the total gaze duration to the surrounding environment, but this dropped to 19% following experience in the environment. Presumably these fixations contribute to the long term memory representation of the environment. Interestingly, the *Experienced* group devoted more gaze time to pedestrians.

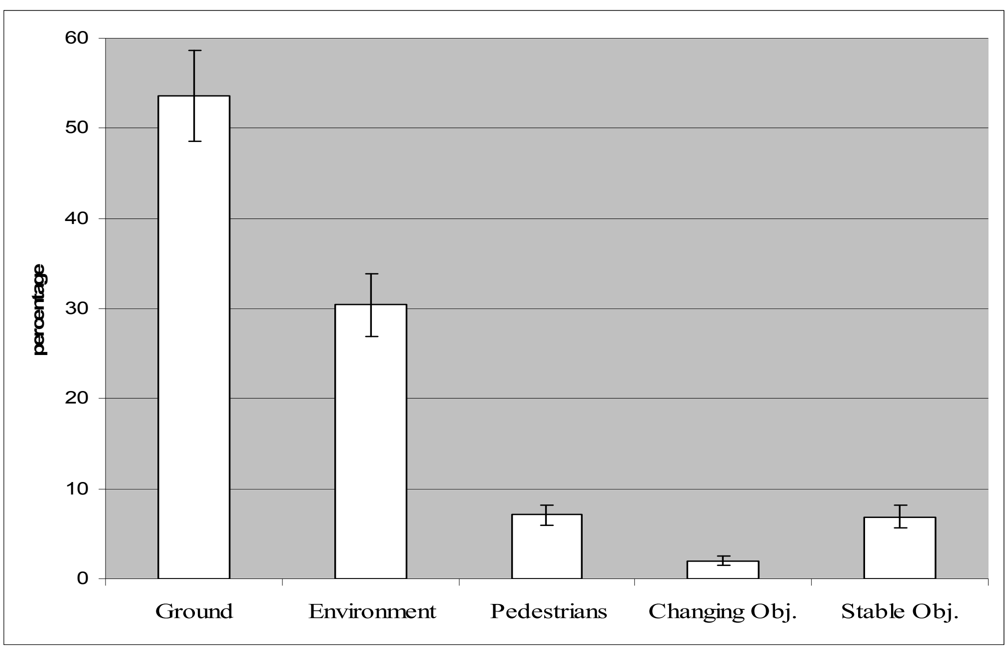


Figure 4-10 Fixation Distribution of the "Inexperienced" group

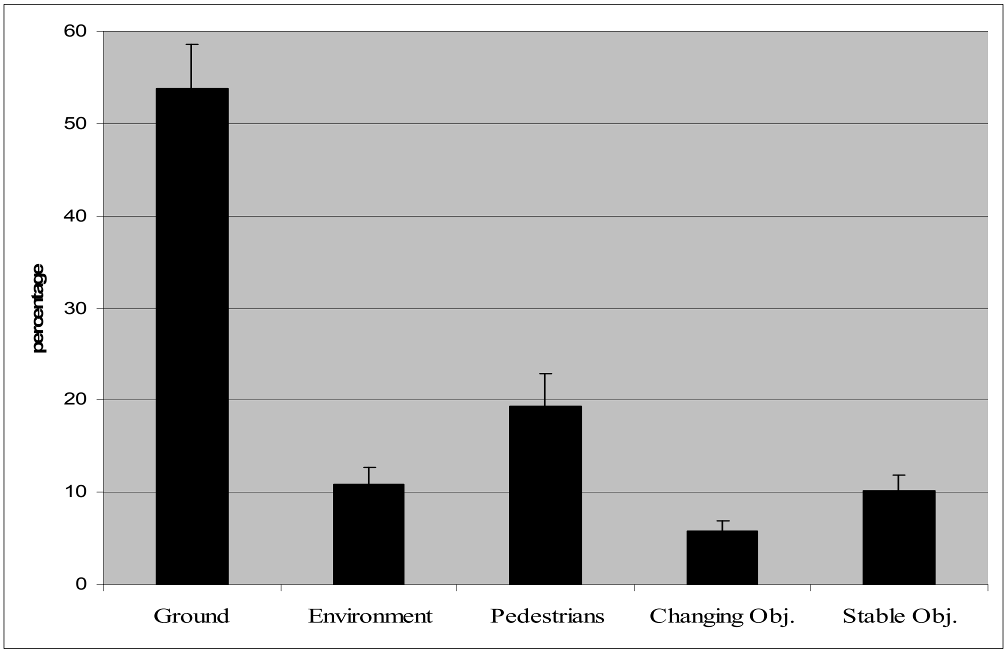


Figure 4-11 Fixation Distribution of the "Experienced" group

The smallest amount of time was spent on the objects in both groups. Note that in these figures, the stable objects have a greater percentage of the gaze distribution because a larger area was occupied by the objects contributing to these figures. When gaze duration is appropriately normalized, the data in Figure 4-9 shows the relative power of the changes to attract gaze.

Correlations between participant reports and fixations

After completing the experiment in the Virtual Environment, participants were asked whether they had noticed any change that had occurred in the environment during the experiment. If their response was positive, they were asked to verbalize their explicit knowledge about those detected changes. In

Table 4.3 you can see the number of participants in each group that detect the different kinds of the changes explicitly.

Table 4-3 Number of the Participants that Detect the Changes Explicitly

Changes / Groups	Object Appeared	Object Moved	Object Replaced	Object Disappeared
Inexperienced	8	4	1	1
Experienced	11	8	3	3

When these verbal reports were compared with their eye fixations, a high correlation between these two variables was observed (Spearman's rho=.586; N=38; $p < .01$).

The correlations of different change types were also investigated. Results showed that the fixation durations on the novel object (Spearman's rho = .397; N=38; $p < .05$), the moved object (Spearman's rho = .520; N=38; $p < .01$) and replaced object (Spearman's rho = .482; N=38; $p < .01$) were highly correlated with the verbal reports while this was not the case for disappearing object (Spearman's rho = .025; N=38; $p = .881$). This suggests that the fixations were accompanied by an awareness of the change.

Effects of computer usage and game playing

A comparable analysis was performed to investigate whether participants' "computer experience" (i.e. how long the participant has been using a computer) and "weekly computer usage" (i.e. how many hours in a week the participant uses a computer) affect their fixation durations on changing objects.

The "computer experience" of the participants varied between 5 and 15 years. The "weekly computer usage", on the other hand, had a larger variation between 5 and 70 hours.

When I took into these variables into consideration as covariates while testing the between-subjects effects (i.e. the differences between the participants according to their experience level in the environment), the results of ANCOVA showed that neither "computer experience" ($F(1,33) = .358$, $p=.554$) nor "weekly computer usage" ($F(1,33)=.583$, $p=.450$) has a significant effect on *Experienced* vs. *Inexperienced* distinction.

The correlation between game playing frequency of the participants and their fixation durations on changing objects was also examined. As mentioned in section 4.1.3, about half of the participants reported that they have never played computer games while others reported they sometimes or frequently play those games. When the correlations were investigated, results showed that the fixation durations had no significant correlation with participants' game playing frequency (Spearman's rho = .037, N=38, $p=.826$).

4.1.8 Discussion

The results of this experiment revealed that familiarity with the environment is an important factor in the distribution of gaze in the environment. In particular, participants spend more time fixating the changed objects if they are familiar with the environment. Participants familiar with the environment fixated changes on average 170 msec longer than participants who were less familiar with the environment.

New and *moved* objects were most effective in attracting gaze, and participants spent approximately 500 msec longer time fixating these objects than stable objects. The similarity of the findings for *moved* and *appeared* objects presumably results from the appearance of the displaced object in a previously unoccupied location. When a *new* object appeared in a previously occupied location, fixation times were much less. Previous studies claimed that abrupt onsets of novel objects attracted a fixation when participants viewed photographic images of natural scenes, regardless of whether they were explicitly told to search for a new object that might appear, or whether less specific memory instructions were given (Theeuwes, 2004; Brockmole & Henderson, 2005a,b). The results of this experiment may also suggest that the appearance of objects in novel locations is indeed an important factor in attracting both attention and gaze in the context on ongoing natural behavior, consistent with these previous studies with non-immersive displays (See Chapter 5 for a detailed discussion of all results that were obtained from each experiment).

The results of this experiment pointed to important findings as discussed above, but additional results were needed for being able to make more clear judgements about the issue. In addition to this need, some previously uncontrolled factors about the design of this experiment, such as the object size, object characteristics and a changing object's possibility of affecting the succeeding change type, were considered as potential problems that would create some bias. Object sizes of the stable and changing objects were different and this circumstance made it harder to make reliable comparisons. Furthermore, different objects changed in different locations in the environment, and this circumstance may also have limited the reliability of the results since the place of change and the physical characteristics of the changing object might be effective on the fixation durations. Finally, in this first experiment, all participants observed all types of changes in the same succeeding order and this might have created a bias because of the possibility that the nature of a detected change type influenced the detectability of its succeeding change type. For example, after a participant observed the appeared object, which is a powerful change according to the results, s/he might fixate on the moved object longer than as it occured alone (i.e. without having a preceeding change).

In order to eliminate these problematic circumstances, the design of the experiment was revised, some pilot tests were applied and the second experiment was conducted on the base of these revisions and the results of pilot applications. In the following sections, you can find the details of the pilot tests and the design of the second experiment.

4.2 Pilot Study for the Experiment 2

In order to examine the differences between four different change conditions as well as the effects of familiarity on these changes, a Desktop Virtual Reality environment was created and pilot experiments were conducted. This virtual environment was then revised and the actual experiments that would provide the required answers were conducted. The eye-tracking technology was again used to follow the position of the observers' eye-gaze, which helps us to measure the attended location in the environment.

The experimental design for this study includes participants exploring a desktop virtual reality environment including both stable and changing objects.

As the environment, a segment of an imaginary town was created by using the Active Worlds software[15]. In the environment, participants followed a rectangular path around a monument (See Figure 4-12).

4.2.1 Virtual Environment

In the environment, in addition to the houses and trees, there were 6 objects, including;

- a park bench
- a street lamp
- a trashcan
- a billboard
- a parasol
- a mailbox

Figure 4-12 A Snapshot from the Environment

[15] For the details of the software visit http://www.activeworlds.com/

Here, all of the objects were stable except for the parasol and the trashcan. These two objects were the changing objects for each experimental group excluding the ones in the new-object condition (See below).

4.2.2 Physical Equipment

The experiments were conducted at the Human-Computer Interaction Research and Application Laboratory at Middle East Technical University. This laboratory is a medium established to design, utilize, and evaluate interactive technologies like web sites and other computer software. The lab consists of an experimentation and a control room (See Figure 4.13). During the experiment, it is possible to get a feedback by recording the facial expressions of the user, hand movements (keyboard & mouse), eye movements, and monitor screen shots.

For this study, Tobii 1750 Eye Tracker was used. The eye tracker is discretely integrated into a 17" TFT monitor without any visible or moving "tracking devices". Participants are allowed to move freely in front of the tracker. This non-intrusiveness ensures that respondents behave naturally, thus providing us with valid data, and allows us to perform long studies without fatigue for the respondent or reduced quality of data. Moreover, it provides a completely natural environment for participants in order to minimize experimental effects.

ClearView analysis software was used for the analysis of the experimental data. This software provides the experimenter with the video of screen contents that are necessary for calculating the fixation durations of the participants.

Figure 4-13 (a) Test Room and (b) Control Room of the Human-Computer Interaction Research and Application Laboratory

4.2.3 Participants

Pilot study participants were 8 undergraduate and graduate students affiliated with the Middle East Technical University. One of the participants was female. The average age of the participants was 26,25. In each experimental group there were 2 participants who were allocated to the groups randomly.

4.2.4 Procedure

The experiment was done in the within-subjects format; four groups were defined according to the type of the change that was visible to the participants. Each group included 2 participants.

In this experimental setting, the distribution of gaze was examined during familiarization trials, and also following changes in the environment. As explained above, the 2 objects in the virtual environment were located at the two sides of the monument which is subjected to change. These changes that were applied to these objects include:

- appearance of a new object
- movement of an existing object
- replacement of an existing object with a new object
- disappearance of an existing object

The participants were divided into four scene change conditions:

- new-object condition
- moved-object condition
- replaced-object condition
- deleted-object condition

In the *new-object condition*, two objects were added to the environment during viewing. In the *moved-object condition*, two existing objects were moved in the environment. In the *replaced-object condition*, two objects were replaced with two different objects. The same objects (a firehydrant & a news-box) in each condition were served as the appearing or the replacing objects. In the *deleted-object condition*, two existing objects were removed from the environment. The disappearing or the moved objects were also same (a parasol & a trashcan) in each condition.

All of the participants were instructed to be familiar with the environment in preparation for a subsequent test in which they would discriminate the learnt objects from pictures in which a detail of an object would be altered (in actuality, this test was never given). No explicit instructions concerning the actual changes were given to the participants.

For each changing condition, there were two changes of the same kind. Each participant was able to see one of the changes after the first turn around the monument (i.e. before he/she becomes familiar with the environment) while the second change occurred after the 6^{th} turn (when we expect the participant to learn the environment). The fixation durations on changing objects were counted for the subsequent 2 turns while ignoring the further fixations (i.e. the fixations on the first changing object were counted only for the 2^{nd} and the 3^{rd} rounds even it will be visible longer).

Each participant was given time to become familiar with the environment while following the same path 8 times. In addition, all of the participants were evaluated individually in order to prevent them to influence each other.

While the participants were exploring the environment, the experimenter watched his/her movements from another monitor in the Control Room. The changes in the environment was applied by using another computer in this room by the experimenter.

After the experiment, each participant was asked to report the changes that they had realized while walking in the virtual environment in order to understand their explicit knowledge about the changes.

4.2.5 Results

The logic of the analysis was calculating the average fixation time of a single object. The total fixation duration was summed up for all stable objects, and also on all changing objects, and then these numbers were divided by the number of turns around the monument in order to get the fixation duration for one round. After this calculation, the result was again divided by the number of objects to obtain the average time fixating a single object in one turn around the monument. For the object that disappeared, fixation on the object's prior location was measured.

During the analysis, all 8 laps in the trial were examined. The first lap, which was the only lap before the changes started to occur, was used only for calculating the average fixation durations for stable objects.

All analyses were done on the video images that were obtained during the experiments. Video records were examined frame by frame to find the fixations. Then, the durations of fixations were summed up to get the numbers that are necessary for the calculations explained above.

Fixation Durations

After examining the data, average fixation duration for the first changing object for all groups was found as 455 msec while it was 1550 msec. for the second changing object. The average fixation duration on stable objects, on the other hand, was 823 msec. The distribution of the fixation durations can be seen from Table 4-4.

Table 4-4 Average Fixation Durations

	First Change	Second Change	Stable Objects
New-object condition *(Participants 6&7)*	820	2060	1195
Replaced-object condition *(Participants 1&4)*	740	1500	796
Deleted-object condition *(Participants 2&5)*	80	1480	704
Moved-object condition *(Participants 3&8)*	180	1160	595

For all conditions, the second change got longer fixations but the difference between the two changes was small when the first change was detected by the participant. In two sessions (Participant 4 & 7), the second change got even shorter fixation durations than the first change (See Figure 4-14).

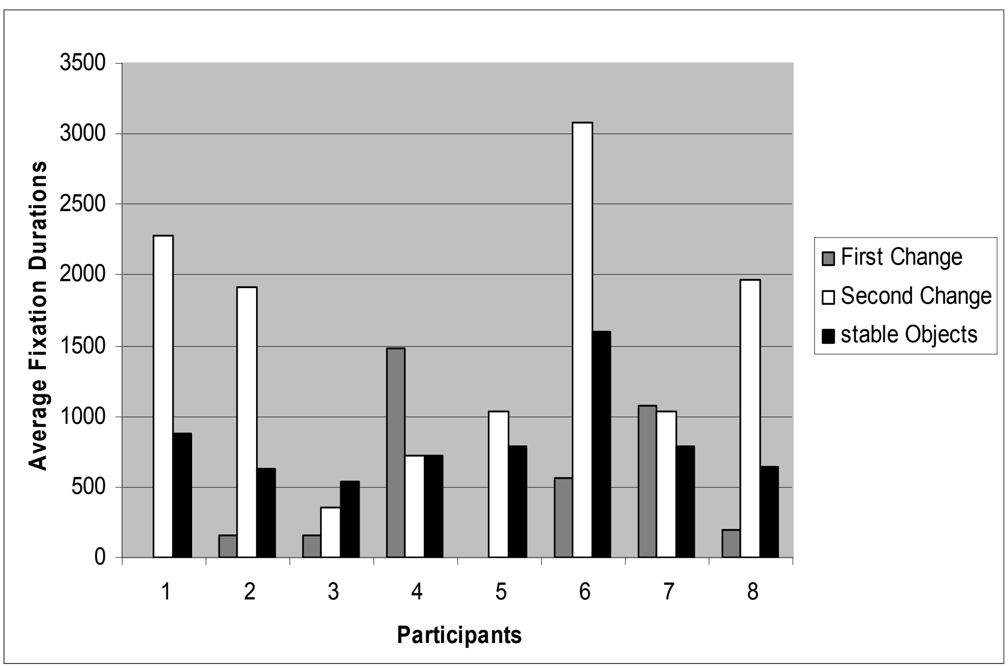


Figure 4-14 Distribution of Fixation Durations According to the Participants

4.2.6 Discussion

After examining the data of this pilot experiment and get some feedback from an expert, some problems that were not previously predicted became visible. The most important problem was the bias that the participants had about the changes during the experiments. Since they saw a change at the beginning of an experimental session they may became aware of the possibility that something would change. Because of this handicap in the design of this pilot experiment, some improvements were made on this design and applied to the design of the second experiment. The details of these improvements and the new experimental design are explained in the next subsection.

4.3 Experiment 2 – Desktop VR

Like the previous design, this experimental design also includes participants exploring a desktop virtual reality environment including both stable and changing objects. As the environment, the segment of an imaginary town, which was created for the pilot study, was used after making some improvements and changes on the objects and procedure (See 4.3.1 & 4.3.4 for detailed information). In the environment, participants again followed a rectangular path around a monument while avoiding pedestrians.

4.3.1 Virtual Environment

In this experiment, the virtual environment was similar to the one that had been used in the first experiment. Like the first experiment there is a rectangular path in which the participants have no chance to see the other three sides while keep going in one side. This issue is important for the changes to be done without the awareness of the participants.

In the environment of the first experiment, there were 8 objects two of which were houses. Since the sizes of the houses were too big compared to the other six objects, the houses were excluded from the measurements in this experimental design.

Figure 4-15 Bird's-eye View of the Environment

In this environment, in addition to the houses and trees, there were 6 objects, including;

- a park bench
- a street lamp
- a trashcan

- a billboard
- a firehydrant
- a mailbox

Here, all of the objects are stable except for the *trashcan*. This object is the changing object for each experimental groups excluding the ones in the new-object conditions (See below).

The stable objects are the same with the ones in the first environment except for the park bench in this environment. There was a water fountain instead of this object in the first design but in order to make the objects in similar sizes, park bench replaced the fountain.

Like in the first experiment, in order to eliminate the random fixations on the objects, avoiding from a potential collision with the pedestrians (See Figure 4.16) is used as a secondary task.

Figure 4-16 A Snapshot from the Environment Showing two of the four Pedestrians

4.3.2 Physical Equipment

The experiments were again being conducted at the Human-Computer Interaction Research and Application Laboratory at Middle East Technical University. Like the pilot experiment, Tobii 1750 Eye Tracker and ClearView analysis software were used for this study.

4.3.3 Participants

Participants were 128 undergraduate and graduate students affiliated with the Middle East Technical University (METU). 85 of them were undergraduate students while 43 participants were graduate students. In each one of the eight experimental groups, there were 16 participants who were allocated to these groups randomly. Of the participants, 75 were female. The gender distribution of the groups can be seen in Table 4.5 below.

Table 4-5 Gender Distribution in the Groups

Group \ Gender	Female	Male	Total
Appear / Experienced	12	4	16
Move / Experienced	9	7	16
Replace / Experienced	9	7	16
Disappear / Experienced	6	10	16
Appear / Inexperienced	6	10	16
Move / Inexperienced	11	5	16
Replace / Inexperienced	11	5	16
Disappear / Inexperienced	11	5	16
Total	75	53	128

The average age was 23,42 (sd = 3,599). The age range was from 18 to 35 (See 17). The academic departments of the participants differred in a range of 34 departments in METU and there were participants from each class level from junior students to PhD candidates.

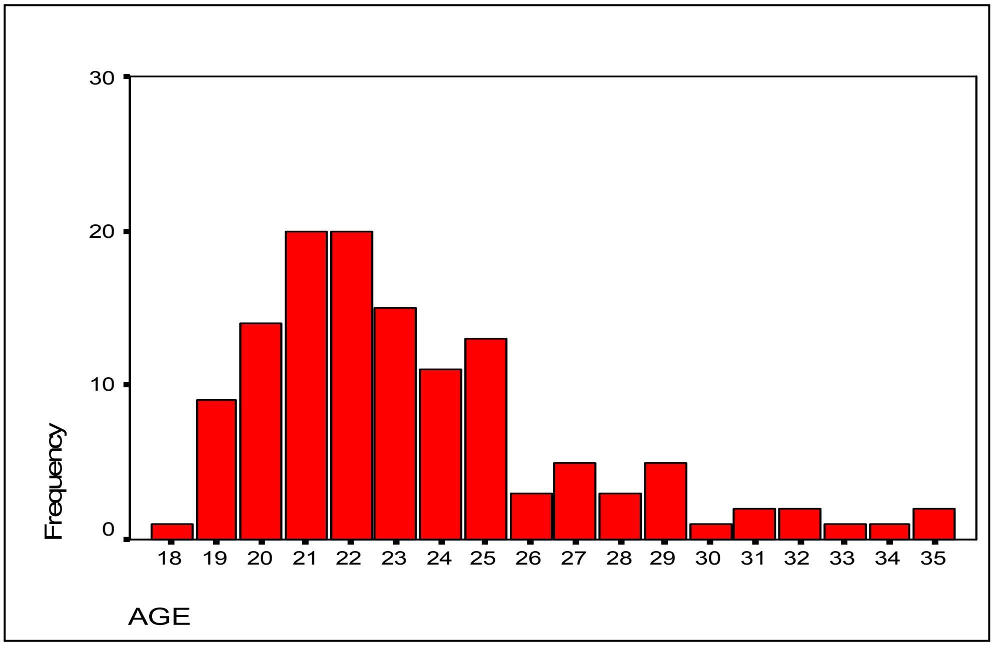


Figure 4-17 - Age Distribution of the Participants

Computer usage habits of the participants were different in a wide range but most of them have been using computers for at least 5 years and almost half of them were playing computer games (See Tables 4.6, 4.7 and 4.8 for the frequencies). By examining the participant reports of the first experiment and the pilot tests, it was observed that general trend about playing computer games was either “never playing” or “frequently playing”, therefore participants of the second experiment were asked only to report whether they play computer games or not. The reason for gathering these personal data was, like in the first experiment, being able to check the potential influences of computer usage habits on perceiving virtual environments. Moreover, all of the participants had normal or corrected-to-normal eyesight.

Table 4-6 Yearly Computer Usage of the Participants

Duration	Frequency	Percentage
Less than 2 years	5	3.9
2-4 years	9	7.0
4-6 years	23	18.0
6-8 years	23	18.0
8-10 years	29	22.7
More than 10 years	39	30.5
Total	128	100.0

Table 4-7 Weekly Computer Usage of the Participants

Duration	Frequency	Percentage
Less than 10 hours	15	11.7
10-15 hours	16	12.5
15-20 hours	12	9.4
20-25 hours	17	13.3
25-30 hours	8	6.3
More than 30 hours	60	46.9
Total	128	100.0

Table 4-8 Game Playing Habits of the Participants

Game Player	Frequency	Percentage
Yes	69	53.9
No	59	46.1
Total	128	100.0

4.3.4 Procedure

The procedure was a combination of the first experiment and the pilot experiment. Same types of changes with the first experiment were applied in a similar environment but this time each participant observed only one type of change while s/he was observing every change type in the first experimental procedure. The first reason for applying this difference in the second procedure was to be able to compare the effects of different change types. Furthermore, I tried to reduce the possibility of a bias that can be created by the detection of a previous change[16]. In other words, if a participant was able to / could detect a change, it became possible for him/her to be aware of the fact that another thing could also be changed in the environment. Therefore, the procedure was revised and between-subjects design was applied to this second experiment.

Similarly with the first experiment, each participant was given time to become familiar with the environment while walking the same path 8 times and this amount of walking is counted as 1 trial. In addition, the experiments were conducted in "one participant at a time" manner, so none of the participants was influenced by the other participants.

[16] Remember that all changes were visible for the participants of the first experiment in a succeeding order; each participant first observed the replacement of an existing object with a new one, then the object disappeared, then appearance of a new object was observed and finally that object moved to another place in the environment.

In this experimental setting, the distribution of gaze was being examined during familiarization trials, and also following changes in the environment. As explained above, one object (a trashcan) in the virtual environment is located at the one side of the monument which is subjected to change. These changes that are applied to this object include:

- appearance of a new object
- disappearance of an existing object
- movement of an existing object
- replacement of an existing object with a new object

In the *new-object condition*, one object is added to the environment during viewing. In the *replaced-object condition*, one object is replaced with another object. The same object (a trashcan or a news-box) in each condition served as the appearing or the replacing object. In the *deleted-object condition*, one existing object is removed from the environment. In the mov*ed-object condition*, one existing object was moved in the environment (See Figure 4-18). The same object (a trashcan or a news-box) in each condition is served as the disappearing or the moved objects.

The changing objects are different for every other person in this experimental setting. In other words, while the trashcan is replaced with a news-box for half of the participants, the news-box is replaced with the trashcan for the other half of the participants. Moreover, while a news-box is the new (i.e. appearing) object and the trashcan is a stable object for the half of the participants, news-box becomes a stable object and the trashcan appears for the other half of the participants. This counterbalancing is applied in order to prevent the possible influences of the property of the objects on fixation durations.

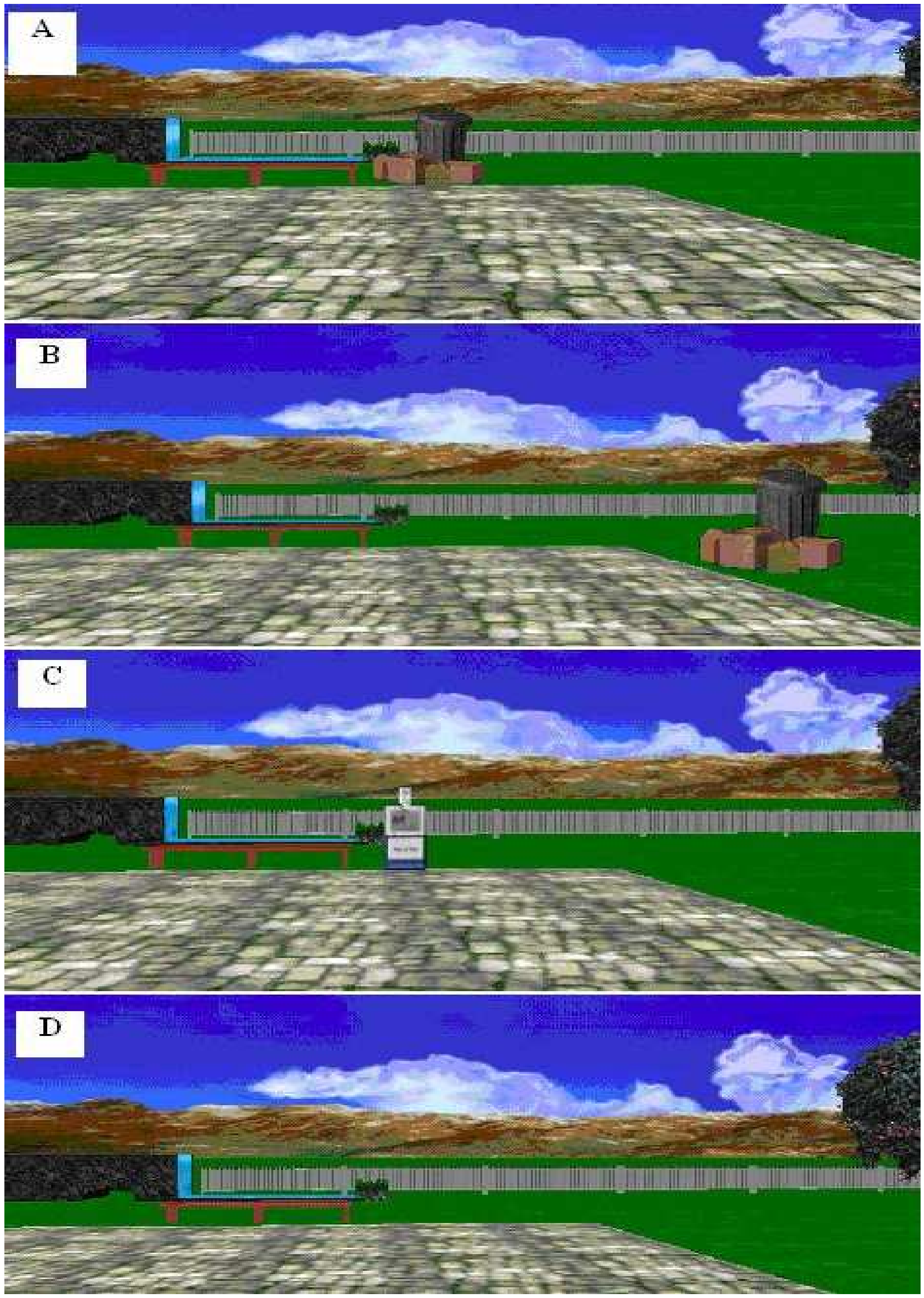
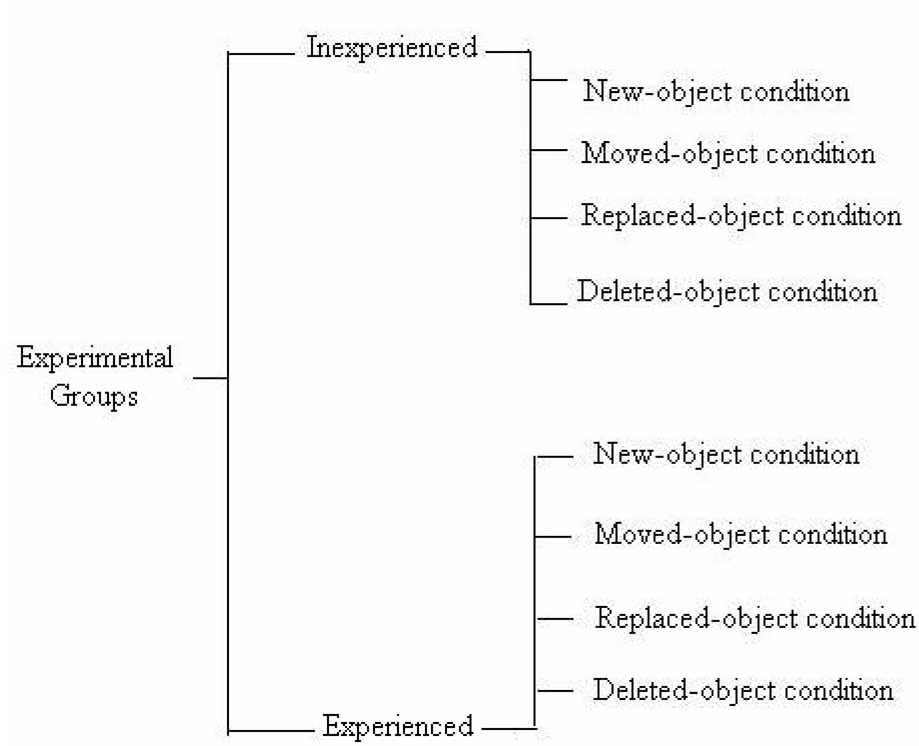


Figure 4-18 A. Object Appeared B. Object Moved C. Object Replaced D. Object Disappeared

As the control condition, average fixation duration on stable objects was used for *new-object, replaced-object* and *moved-object* conditions. For the deleted-object condition, on the other hand, average fixation duration on the empty spot of the deleted object was used as the control condition. This gives us the possibility of fixating to the empty spot without any effect of change. This value was obtained by analyzing the fixation durations on the specific empty spot before the appearance of the object in the new-object condition.

As explained earlier, the experiment was designed to be in the between-subjects format; eight groups were defined according to the participants' familiarity with the experimental environment and the type of change they see. The groups are divided as follows:



Each group included 16 participants. The first four groups are named "Inexperienced" since they were able to see the changes after the first turn around the monument (i.e. before they become familiar with the environment). On the other hand, the second four groups were more familiar with the environment since they had 6 turns around the monument before the changes occurs. Because of this reason, they are named as "Experienced". The participants were sequentially allocated to one of these eight groups.

The fixation durations on changing objects were counted for the subsequent 2 turns while ignoring the further fixations (i.e. the fixations of the "Inexperienced" groups were counted only for the 2^{nd} and the 3^{rd} rounds even though the changing object will be visible longer). The reason for ignoring further fixations for the "Inexperienced" participants was being able to compare the results more reliably since "Experienced" participants could observe the changes only for the last two turns (i.e. the 7^{th} and the 8^{th} rounds).

All of the participants were instructed to be familiar with the environment while avoiding pedestrians. No explicit instructions concerning the actual changes were given to the participants.

Total duration of one session was varied between 10 and 15 minutes according to the computer usage of the participant. After each session of the experiment, the participant was asked to report the changes that they had realized while walking in the virtual environment in order to understand their explicit knowledge about the changes.

4.3.5 Hypotheses

The results of this experiment test the following hypotheses:

H_1: Fixation durations will be longer for the *Experienced* groups for each of the four different changing types.

H_2: Fixation durations will be less in the Desktop VR environment as compared to the Immersive one in the first experiment.

4.3.6 Results

The logic of the analysis was calculating the average fixation time of a single object. The gaze durations were counted by analyzing the video results of the experiments frame by frame. Fixations were defined as a constant location within a 1 deg radius for a period of 100 msec or more. I was interested in the total amount of time fixating a particular object. This might be composed of a single long fixation or several shorter fixations on the object. I will refer to this as gaze duration to avoid confusion with the duration of a single fixation.

The location of eye fixation was taken as an index of the locus of attention within the scene. The total duration of all fixations on stable objects was found, and also on all changing objects, separated by category of change. Gaze was also divided up into fixations on the background and pedestrians. The total fixation duration was summed up for all stable objects, and also on all changing objects, and then these values were divided by the number of circuits around the monument, and the number of objects to give the total time fixating a single object in one circuit around the monument, averaged over objects. For *deleted-object* condition, fixation on the object's prior location was measured.

All analyses done on the video files that were recorded during the experiments. These videos then converted to a format that can be usable for PCs. After this convertion, the images were imported to the Windows Movie Maker Application and then obtained video records were examined frame by frame to find the fixations. Then, the duration of each fixation was calculated in milliseconds. After these analyses, the durations of fixations were summed up to get the numbers that are necessary for the calculations explained above.

Effects of Familiarity

After examining the data, average fixation duration on the changing object for the "Inexperienced" groups was found as 709 msec while that of the "Experienced" groups was 1922 msec. In addition, average fixation duration for all stable objects for the "Inexperienced" groups was 848 ms and that of the "Experienced" groups was 867 msec (See Figure 4-19).

For the statistical analysis, 2 (Stable, Changed) * 2 (Experienced, Inexperienced) * 3(new, moved, replaced) ANOVA design applied and the results showed a significant difference for fixation durations on the changed (new, moved, replaced) object between the two groups ($F (1, 90) = 22.102$, $p<.01$), while almost no difference was observed between the two groups for their fixation durations on stable objects ($F (1, 90) = 1.410$, $p = .238$).

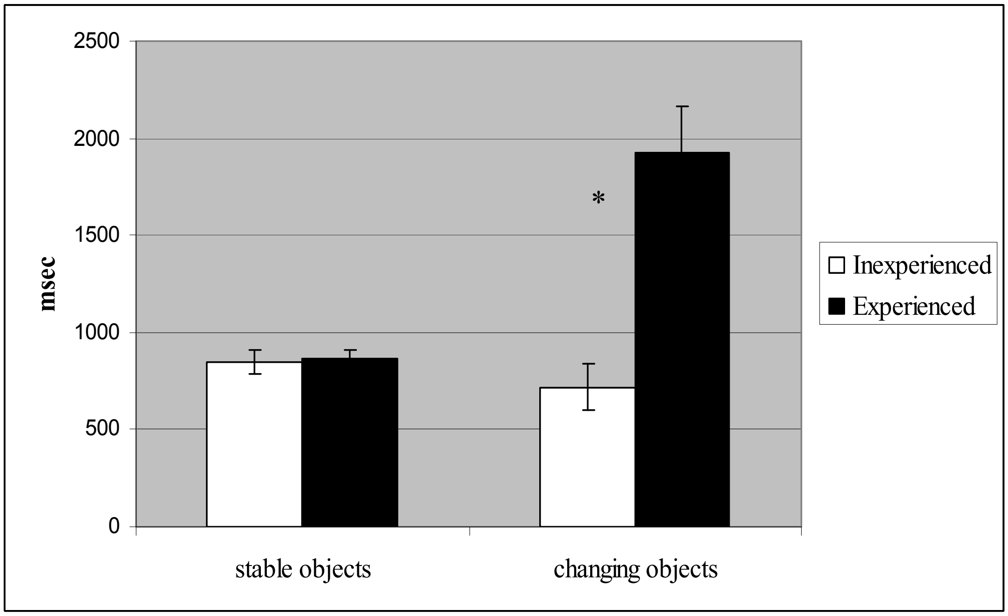


Figure 4-19 Average Fixation Durations

Since the fixation durations for the deleted object fell on an empty spot rather than an object as in the other change types, effects on this change type examined separately in a 2(Stable, Deleted) * 2 (Experienced, Inexperienced) ANOVA design. The result of this analysis also showed that there is a highly significant difference between the two groups according to their fixation durations for the deleted object ($F(1, 30) = 12.138$, $p < .01$). On the other hand, when the fixation durations of the participants, who examined the virtual environment in the deleted-object category, were examined, again no significant effect of familiarity (*Experienced* vs. *Inexperienced*) was observed for the stable objects ($F(1, 30) = 1.665$, $p = .207$).

Both groups had almost identical fixation durations on the stable objects. However, participants who had experience in the environment looked substantially longer at the changing objects, an increase of approximately 1200 msec. This finding, like the one in the first experiment, suggests that prior experience in the environment indeed increased the likelihood that participants would look at a change in the scene.

Effects of Change Type

There were four types of changes that occured in the environment during the experiments. The effects of these different changes were also statistically examined. These data are shown in Figure 4-20 for each change type and in Figure 4-21 for the two bunch of groups. These changes and their representations (in italics) in the figure are listed below:

- appearance of a new object *(New)*
- movement of an existing object *(Moved)*
- replacement of an existing object with a new object *(Replaced)*
- disappearance of an existing object *(Deleted)*

When the fixation durations on stable and changing objects were compared, without considering the effect of familiarity, average fixation duration on the changed object of each type were 1948 msec, 1202 msec, 1593 msec and 230 msec for *new-object, moved-object, replaced-object* and *deleted-object* conditions respectively (See Figure 4-20).

Average fixation durations on stable objects were not much fluctuating like that of changing ones. These values were found to be almost equal for all change types. It was 836 msec for *new-object condition*, 819 msec for moved object condition and finally 844 msec for replaced object condition. As mentioned before, for *deleted-object condition*, the possibility of fixating on the empty spot of the deleted object is used as control condition instead of stable objects. This value was found to be 149 msec.

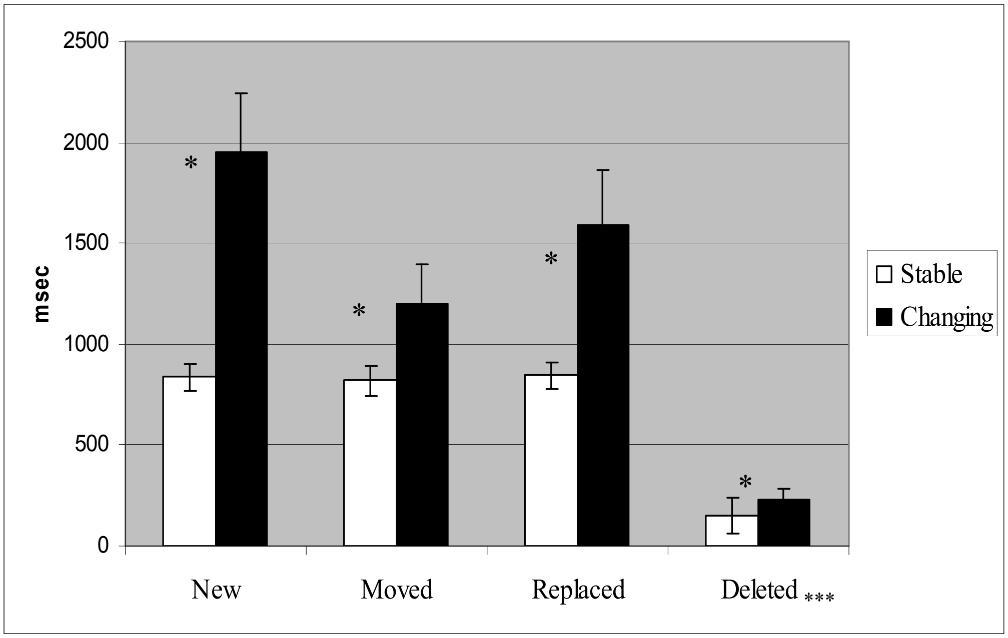


Figure 4-20 Average Fixation Durations on Stable & Changing Objects for Different Change Types (* For Deleted-object condition, the possibility of fixating on the empty spot of the deleted object is used instead of stable object)**

When this data was statistically examined, the results showed that the fixation durations on the changing (new, moved, replaced) objects were significantly different from that of stable objects ($F(1, 90) = 29.575$, $p < .01$). Then, the relationship between this difference and the three types of change was examined and the results showed no significant interaction ($F(2, 90) = .761$, $p = .470$). In other words, no significant difference was observed between *new, moved* and *replaced-object* conditions on the level of the changed condition within the factor “change vs. stable”.

For deleted-object condition, the possibility of fixating on the empty spot of the deleted object was used instead of stable object as explained above. The results of this comparison showed that the difference between participants' fixation durations on the empty spot of the *deleted-object* and the possibility of fixating on that empty spot was also significant ($t(31)=4.665$, $p < .01$).

In other words, the average fixation durations on each type of changed object were not significantly different from each other but all change types showed a significant difference from those fixation durations on their respective baselines for control conditions.

Effects of Familiarity Combining with the Change Type

When the effects of familiarity and the type of change were combined, it was found that "Experienced" group fixated longer on every type of the changing object than the "Inexperienced" group (for statistical results see "Effects of Familiarity" section).

Numerical data showed that *New* object condition was the most powerful change type to attract the gaze of the participants in both groups. In this manner, *Replaced, Moved* and *Deleted* object conditions followed the *New* object condition in a descending way while the *Deleted* object condition was the least powerful change type.

The biggest difference between the average durations of fixations of the two groups was for the *New* object (1815 msec). Furthermore, the differences between the average durations of fixations of the two groups on *Replaced,* Moved and *Deleted* objects were 1506 msec, 709 msec and 294 msec, respectively (See Figure 4-21).

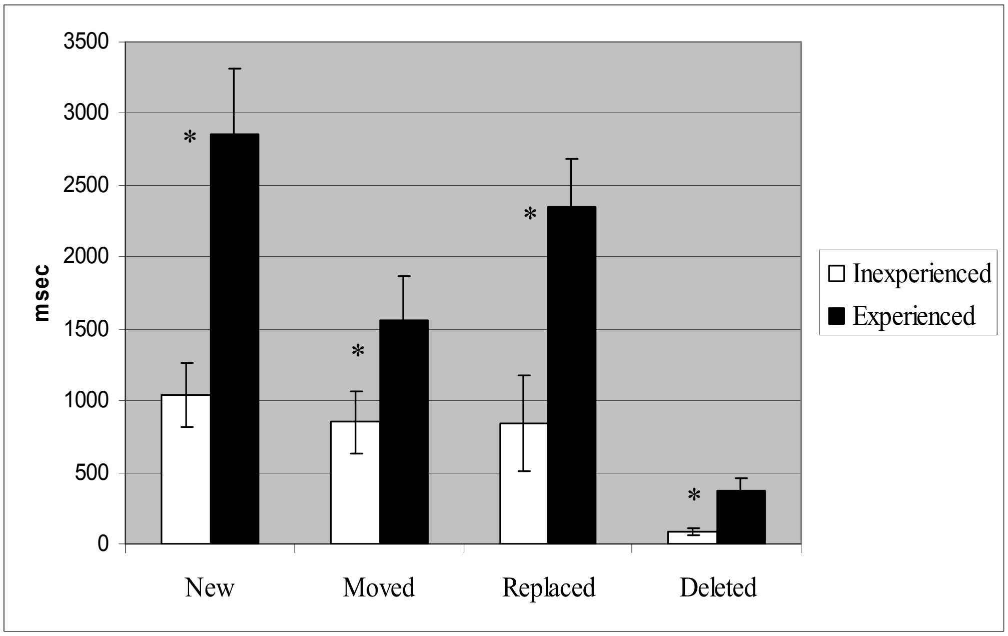


Figure 4-21 Average Fixation Durations for Different Change Types

In order to understand the relationships between these different changes a Scheffe procedure was applied as a further analysis. This analysis provided us with some findings which showed the 95 % confidence interval boundaries for deciding whether the differences between the average fixation durations in new, moved and replaced object conditions were significant or not. After this calculation, the difference between the means of new and moved objects (746 msec) exceeded 670 msec, which was necessary for significance according to the Scheffe procedure.

Time Duration Before the First Fixation on Changed Object

In this experiment, another variable was also examined. This variable was the time duration before the first fixation on the changed object after it became visible to the participant. Here, the aim was further investigating the effects of familiarity and change types on the nature of the detection process. These durations were calculated again by analyzing the videos in a frame-by-frame manner. First, the time of change (i.e the time that the change became visible to the participant) is noted, then the time of first fixation on that changed object is noted and the duration between these two time spots was calculated for the analysis.

When the data were examined, it is found that there is a significant difference between the *Experienced* and *Inexperienced* groups ($F(1,112) = 29.107$, $p<.01$) about their time durations before the first fixation on the changed object. The average duration for the Experienced group was 9689 msec while that of Inexperienced group was 36375 msec. In other words, the duration was almost four times longer for the Inexperienced participants.

When the results of the t-tests were examined, it was found that these differences between the *Experienced* and *Inexperienced* groups were significant for all change types except for the *new-object* condition which has a difference value that is close to significance ($t(28)=-1.994$, $p=.056$). Experienced participants' time durations before the first fixation on the changed object were significantly different from those of Inexperienced participants for the *moved-object* ($t(28)=-3.368$, $p<.01$), *replaced-object* ($t(29)=-3.534$, $p < .01$) and *deleted-object* ($t(21) = -4.551$, $p < .01$) conditions.

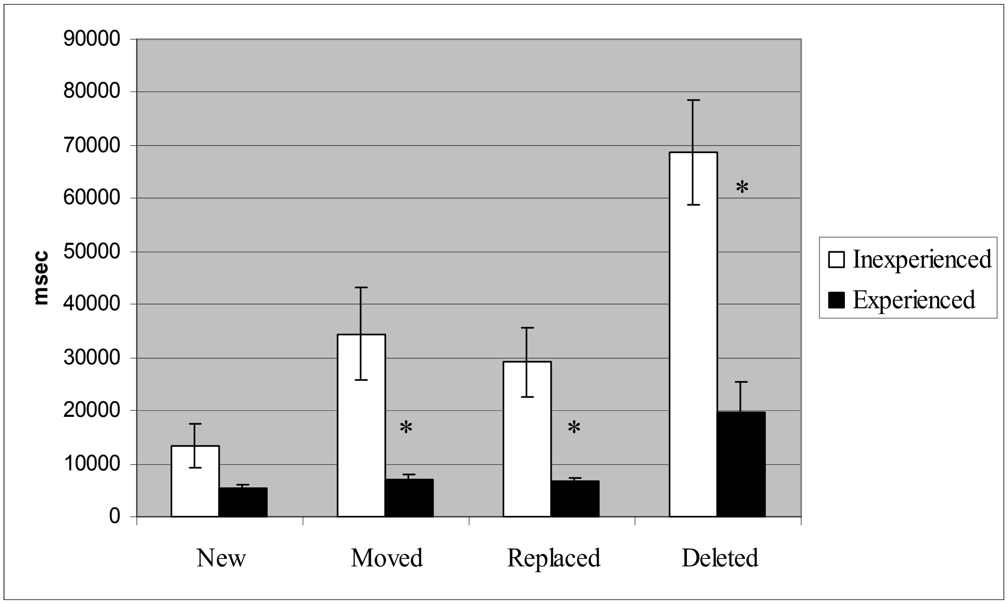


Figure 4-22 – Time Durations before the first fixation on the changed object after it became visible to the participant

When the durations for these two types of groups were seperately examined, the data revealed that the type of change has a significant effect on both *Experienced* ($F(3,59) = 5.345$, $p<.01$) and the *Inexperienced* ($F(3,47) = 7.968$, $p<.01$) type of groups.

Within the different change conditions, the shortest time duration was observed for the *new-object* conditions of the *Experienced* (5405 msec) and *Inexperienced* (13241 msec) groups. In this row, the second place was occupied by the *replaced-object* conditions of the *Experienced* (6618 msec) and *Inexperienced* (29105 msec) groups. *Moved-object* condition was not that much effective. For that change type, the durations were 7004 msec and 34446 msec for *Experienced* and *Inexperienced* groups respectively. Finally, the longest time durations were observed for the *deleted-object* conditions of the *Experienced* (19728 msec) and *Inexperienced* (68709 msec) groups.

Individual Differences

75 of the participants were female and all participants were in a range of ages between 18 and 35. Even though the study did not aim to explore the correctness of a hypothesis that includes individual differences, as an additional finding, gender and age differences were also examined for this second experiment.

First, t-test was used to examine the effect of gender on participants' performance and no significant difference was found between males and females according to their fixation durations on changing objects ($t(126)= 0.656$, $p = .513$). Moreover, the correlations between the ages of the participants and their fixation durations on changing objects were examined, and again no correlation was found (Pearson's $r = .103$; $N=128$; $p= .246$). Finally, departments of the participants were categorized as natural sciences or social sciences and the correlations between these categories with the participants' fixation durations on changing objects were examined, and again no correlation was found (Pearson's $r = -.063$; $N=128$; $p= .478$).

Gaze Distribution

I also examined how participants distribute their gaze in the environment. The location of the fixations was classified into fixations on the path, surrounding environment (for example the grass, the monument, or in the distance), pedestrians, or objects, either changing or stable. The proportion of time spent fixating on each of these regions are plotted in Figures 4-23 & 4-24 for the *Inexperienced* and *Experienced* groups respectively.

Most of the fixations were located on the walking path for both Inexperienced and Experienced groups (56% and 53% respectively). This may reflect the ongoing demands of walking and staying on the path.The *Inexperienced* group devoted 28% of the total gaze duration to the surrounding environment, and it was almost the same for the *Experienced* group with a percentage of 26%.

The smallest amount of time was spent on the objects and pedestrians in both type of groups. The percentage of the fixations on stable objects (9% and 10% for the *Inexperienced* and *Experienced* groups respectively) and pedestrians (6% and 7% for the *Inexperienced* and *Experienced* groups respectively) were also close to equal for the two type of groups. The only difference on the gaze distribution of the two type of groups was for the changed objects. The percentage of the fixation on the changed object was 1% for the *Inexperienced* groups while that of *Experienced* groups was 4 %.

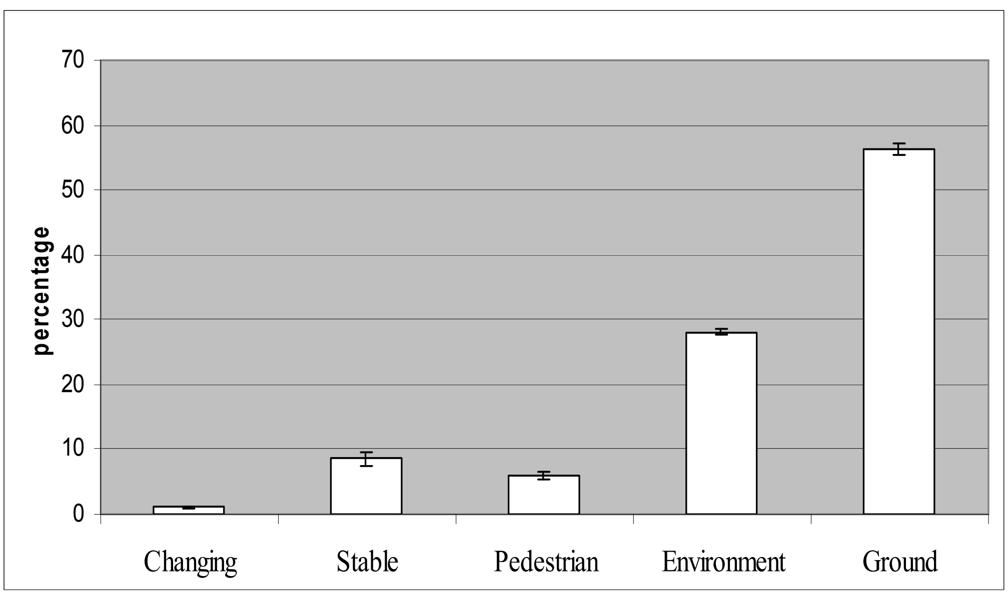


Figure 4-23 Fixation Distribution of the "Inexperienced" group

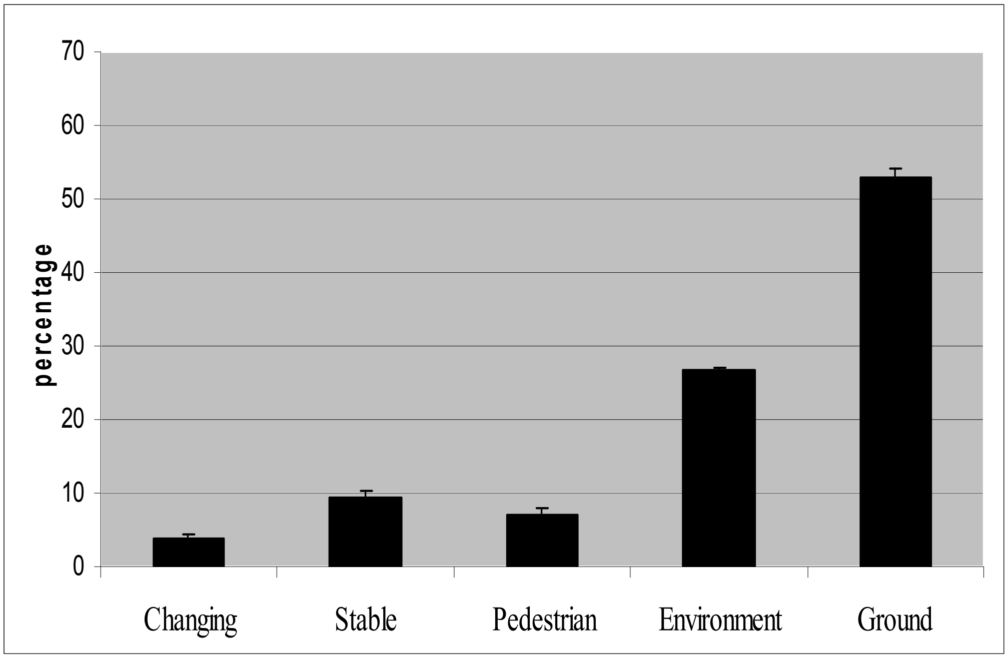


Figure 4-24 Fixation Distribution of the "Experienced" group

Correlations between participant reports and fixations

Like the first experiment, after completing the experiment in the Virtual Environment, participants were asked whether they had noticed any change that had occurred in the environment during the experiment. If their response was positive, they were asked to verbalize their explicit knowledge about those detected changes.

In Table 4.9 you can see the number of participants in each group that detected the different kinds of changes explicitly.

Table 4-9 Number of the Participants that Detected the Changes Explicitly

Changes / Groups	Object Appeared	Object Moved	Object Replaced	Object Disappeared
Inexperienced	6	2	3	0
Experienced	11	9	11	5

When we compared these verbal reports with their eye fixations we found that there was a high correlation between two (Spearman's rho = .422; N=128; $p<.01$). The average fixation durations on changing objects were longer for the participants who reported that s/he had recognized the change. This suggests that the fixations were accompanied by an awareness of the change. Therefore, either participants looked at the changing area because they noticed that something had changed at that specific spot in the environment or their longer fixation durations on the changing objects provided them with the awareness of that specific change.

Effects of computer usage and game playing

A comparable analysis was performed in order to investigate whether participants' "computer experience" (i.e. how long the participant has been using a computer) and "weekly computer usage" (i.e. how many hours in a week the participant uses a computer) affect their fixation durations on changing objects. The "computer experience" of the participants varied between 1 and 15 years. The "weekly computer usage", on the other hand, had a larger variation between 5 and 70 hours.

When these variables were taken into consideration as covariates; ANCOVA results showed that "computer experience" ($F(4,123) = .035$, $p=.851$) did not make a difference but as an interesting and a different result from the first experiment "weekly computer usage" ($F(4,123)=5.299$, $p<.05$) showed a significant effect.

The correlation between game playing frequency of the participants and their fixation durations on changing objects was also examined. As mentioned in the section 4.3.3, about half of the participants reported that they have never played computer games while others reported they sometimes or frequently play those games. When the correlations were investigated, results showed that the fixation durations had no significant correlation with participants' game playing frequency (Spearman's rho = .024, N=128, $p=.393$).

4.3.7 Discussion

The results of this experiment again revealed that familiarity with the environment is an important factor in the distribution of gaze in the environment. In particular, participants spend more time fixating the changed objects if they are familiar with the environment. The participants who were familiar with the environment fixated changes on average 1200 msec longer than the participants who were less familiar with the environment. Time duration for the first fixation to be attracted to the changed object after it became visible was another indicator of the effect of familiarity on change perception. *Experienced* participants looked at the change much more quicker than *Inexperienced* ones.

In this experiment the effects of familiarity and the type of change were combined, and it was found that “Experienced” group fixated longer on every type of the changing object than the “Inexperienced” group. When we look at the numerical data about the fixation durations, *New* object condition was observed as the most powerful change type to attract the gaze of the participants in both groups. In this manner, *Replaced, Moved* and *Deleted* object conditions followed the *New* object condition in a descending way while the *Deleted* object condition was the least powerful change type like it was in the first experiment. Despite these numerical differences, statistical analysis showed that the average fixation durations on each type of changed object were not significantly different from each other while all change types showed a significant difference from those fixation durations on their respective baselines for control conditions. This result claims that the only important difference for change detection was participants’ familiarity level without regarding what type of change they observed.

Even if the type of change seems to have no interaction with change detection, it has actually a significant effect on both *Experienced* and the *Inexperienced* type of groups when we look at the average time durations for the first fixation to be allocated on the changed object after it becomes visible to the observer. It means familiarity influences change detection in every type of change even if the nature of these changes are indeed different from each other. This finding also shows the power of learning (i.e. familiarity in our sense) since it has its effects on each different change type regardless of its nature.

It is also important to note that these changes (i.e. appereance, disappearance, replacement & movement) occurred while the objects were out of the field of view of the participant, and so were not accompanied by a retinal transient.

Thus their attentional prioritization must be a consequence of the difference between the current image and the participant's stored memory representation of the scene. This is consistent with Brockmole and Henderson's (2005a) result showing that changes occuring during a saccade had the power to attract fixations when participants had the opportunity to construct a memory representation of the scene during a prior 15 sec exposure to the image (See Chapter 5 for a detailed discussion of the results).

CHAPTER 5

DISCUSSION & CONCLUSION

The results of the two experiments revealed that familiarity with the environment is an important factor in the distribution of gaze in the environment. There were also other common and different findings of these experiments. In this chapter, the study and its findings are summarized and different aspects of the results are discussed in more detail. Then, conclusions are drawn on the basis of these discussions while spelling out the limitations of the study and the posibilities for future research.

5.1 Summary of the Study

This study was designed to examine the differences in visual attention under different levels of familiarity in a virtually modeled natural environment. For this purpose, an Immersive and a Desktop Virtual Reality Environments were created and experiments that would provide the required answers were conducted in these environments.

The eye-tracking technology was also used to follow the position of the observers' eye-gaze, which helped us to measure the attended location in the environment.

The first experiment was conducted in an immersive virtual environment where 38 participants were given time to become familiar with the environment while walking the same path 6 times. Distribution of gaze was examined during familiarization trials, and also following changes in the environment, including appearance, disappearance, movement, or replacement of an object. The results of this experiment revealed that familiarity with the environment is an important factor in the distribution of gaze in the environment (Karacan & Hayhoe, in press). On the other hand, this first experiment did not give a deep insight about the differences between the four change types.

After examining the first bunch of data, I needed additional evidence about the nature of the different change types and started to design a new experiment for verifying and getting further answers. Before starting to conduct the second experiment, a pilot experiment was conducted to see possible difficulties and biases in a potential design. This experiment was conducted in the within-subjects format; four groups were defined according to the type of the change that was visible to the participants. Each group included 2 participants making a total of 8.

After examining the data of this pilot experiment and getting some feedback from an expert, some problems that were not previously predicted became visible. The most important problem was the bias that the participants had about the changes during the experiments. Because of this handicap in the design of this pilot experiment, some improvements were made on this design and it was applied for the second experiment. In order to examine the differences between four different change conditions as well as the effects of familiarity on these changes, the virtual environment that was created for the pilot experiment was revised and the second experiment was conducted in this environment.

The procedure was a combination of the first experiment and the pilot experiment. The experiment was designed to be in the between-subjects format; eight groups defined according to the participants' familiarity with the experimental environment and the type of change they see.

Table 5-1 A Brief Summary of the Experiments

	Experiment # 1	**Experiment # 2**
Main Question	If subjects become familiar with an environment, are changes more likely to attract attention?	What are the relationships between different change types and effects of familiarity on these changes?
Participants	within-subjects 38 participants	between-subjects 128 participants
Environment	Immersive VR	Desktop VR
Main Finding	Participants are more likely to fixate changed objects if they are familiar with the environment.	Familiarity is an important factor for detecting changes and all change types attract attention while novelty is the most detectable change type

Results of this second experiment revealed that the average fixation durations on each type of changed object were different from not only each other but also from those fixation durations on their respective baselines for control conditions. The *new-object* condition was observed as the most powerful change type to attract the gaze of the participants while the *deleted-object* condition was the least powerful change type in both experiments. In addition, the overall results (i.e. the results of both experiments) showed that participants' familiarity with 3D virtual environments significantly increases the probability that gaze will be attracted to changes in the scene (for a brief summary of the two experiments see Table 5.1)

5.2 Discussion

In this subsection the results of the experiments are discussed in detail in order to understand what those numerical data tells us about the nature of change detection in virtually modeled natural environments. For this issue, first the two experiments are compared on the basis of the findings and then all results are discussed according to their contributions on the previous findings about attention controlling mechanisms and change blindness phenomena while also spelling out the puzzling aspects of present findings.

5.2.1 Comparison Between Experiments 1 & 2

The results of the second experiment were almost identical with those that were obtained from the first one. When we look at the Figures 4-18 and 4-19, we see the same pattern. The average fixation durations on changing objects were much more longer for the participants who were "experienced" than that of "inexperienced" participants. On the other hand, almost no difference was observed between the fixation durations on stable objects. Thus, data of both experiments revealed that familiarity is an important factor for change detection.

When the change types were seperately examined, again a similar picture between the two experiments was observed. The pattern of the results were parallel (See Figure 4-9 (p.72) and 4-21 (p.105)). In addition, statistical analysis revealed that neither of the results of two experiments showed a significant difference between *new, moved* and *replaced-object* conditions. This result claims that the only important difference for change detection was participants' familiarity level without regarding what type of change they observed.

Previous studies introduced a debate about the detectability levels of appeared and deleted objects. Some studies claimed that appearance of an object might be more effective than its disappearance (Boot, Kramer & Peterson, 2005; Brockmole & Henderson, 2005). On the other hand, it was also claimed that identification of object deletion is more likely than the identification of object addition (Mondy & Coltheart, 2000). Both experiments of this study favored the former explanation since the *new-object* condition was observed as the most fixation gathering change type to attract the gaze of the participants while the *deleted-object* condition was observed as the least powerful change type in this manner. This finding was strengthened by the analysis of the time duration before the first fixation on the changed object. In that analysis, the average time duration for the first fixation on the *new-object* after it became visible to the participant (5405 msec and 13241 msec for Experienced and Inexperienced groups respectively) was much more less than that of *deleted-object* (19728 msec and 68709 msec for Experienced and Inexperienced groups respectively). This result shows that detecting that something new has appeared in the environment is much faster than detecting that something old has disappeared.

Fixation distributions were also similar for the two experiments which indicate that the divisions of the two virtual environments were similarly interesting and informative for the participants. The fixation durations on the ground were always the longest and this may reflect the participants' ongoing demands for staying in the path. This may also indicate a similarity about participants' movement/walking styles between the immersive and desktop virtual reality environments and their common demands with moving in real world places. On the other hand, percentage of fixation durations on pedestrians dropped in the desktop VR environment and this may because of the difference between passively viewing on screen and actual walking. Therefore, if participants were able to think that their body was a part of the environment, that feeling changed their perception. This finding points out an important cognitive factor that the immersion (i.e. feeling of presence) is an important factor on visual perception due to user's physical demands in the environment. Here, it should be noted that, even though no presence questionnaire was used in the experiments of this study, we are claiming immersion as being higher in the virtual environment that was used for the first experiment since it was an immersive VRE.

Individual differences (i.e. age, gender & academic background) were found to be not correlated with the performance of the participants in the tasks of both experiments.

After analyzing the effects of computer usage habits, neither participants' computer experience (i.e. for how many years they have been using computers) nor their game playing frequencies showed a significant correlation. Weekly computer usage, on the other hand, showed a significant effect on the change detection performances of the participants in the second experiment while it was not an important factor in the first experiment. This may be due to the similarity between the usages of desktop VR technology with desktop computers. Moreover, there is a possibility that the immersive VR environments could be felt as more naturalistic environments than desktop ones. If so, observing no significant effect of participants' game playing habits on their performance in an immersive environment becomes reasonable since less logical correlation can be set between game playing and natural behavior than that of between game playing and computerized behavior.

The correlations between participants' performance and their explicit reports were high for both experiments. The average fixation durations on changing objects were longer for the participants who reported that s/he had recognized the change. This suggests that the fixations were accompanied by an awareness of the change rather than random gaze movements.

The second experiment not only strengthened the results of the first experiment but also added new findings about the time spent for the first fixation was attracted to the change and the possible threshold levels in order for familiarity to change into recollection. While the former again revealed the effect of familiarity on change perception, the latter gave us a clue about the necessary fixation duration for something to be recognized explicitly.

In summary, no major difference observed between the results of the two experiments. Both experiments gave parallel answers to the questions of this study and the most important finding was the high correlation between participants' familiarity level and their change perception performance.

5.2.2 Attention Controlling Factors

The present results point to an important factor in controlling attention in the context of natural behavior. Most environments are familiar. Even environments that are novel will contain many statistical similarites with environments that have been experienced in the past (for example street scenes, classrooms, stairways and halls). Thus a mechanism which draws attention to regions that do not fit stored memory representations will be generally useful. What are the mechanisms by which this attraction occurs? Some insight may be given by models of early cortical processing where feedback signals from higher cortical areas carry a model-based prediction to lower areas (Rao & Ballard, 1999). When there is a mis-match between the input and top-down predictive signals, the residual signal is transmitted to higher cortical areas and generates a revised prediction. If the predictive signal is based on a stored memory representation of the scene, scene changes may generate a mis-match, or residual signal, that prompts a re-evaluation of the scene, and may thereby attract attention. This kind of a mechanism seems likely to be more robust than saliency based models such as Itti & Koch (2001), where intrinsic stimulus properties are hypothesized to attract attention. Such models are inflexible, and do not account for more than a small proportion of the fixations observed in natural behavior (Jovanevic et al, 2006; Hayhoe & Ballard, 2005; Rothkopf et al, 2005). A mis-match mechanism such as that indicated here, can do the job of saliency-based models by allowing attentional deployment to regions that are not part of the ongoing task, and where something unexpected might occur. Such a mechanism is not easily classified as top-down or bottom-up, and can serve as a useful adjunct to attentional deployment that is governed by the ongoing task. It can neither be classified as top-down since it is stimulus driven (i.e. attention is driven by environmental events) nor it can be classified as bottom-up because it is also goal-driven (i.e. ongoing task demands and prior knowledge are important). This is of course speculative, but the robustness of the effects observed in the current experiments suggests that a model of this general kind is required.

This study also revealed another important factor about the control of attention which is “learning”. As it was stated in the first chapter, one of the aims of this study was understanding whether learning is important for attention or only the reverse is true. When we look at the distribution of gaze in the virtual environments of this study, an effect of real-world knowledge is observed such that pedestrians did not get too much fixations even if avoiding from a potential collision with them was one of the main tasks for the participants. This may due to the fact that in our normal, daily lives we do not bump into pedestrians and since the participants have learned this fact, like all other people in the world, they allocated their attention accordingly. Furthermore, when they became familiar with the environment (i.e. they learned the environment somehow) their fixation durations on changing objects were increased. This may also because of the fact that they learned where to allocate their attention during the trials. Putting all together we can claim that not only attention is important for learning but also learning is important for attention.

5.2.3 Contributions to Change Blindness

The results also have significance for understanding change blindness. Poor performance in change blindness experiments has typically been interpreted as evidence for limited scene memory (Rensink, 2000; Simons, 2000a; O’Regan, 1999). Change blindness clearly reflects the limitations of short term memory, although it is now clear that visual long term memory for scenes is quite extensive (Hollingworth & Henderson, 2002; Melcher & Kowler, 2001;

Melcher, 2005). Why, then, have the demonstrations of change blindness been so effective?

The present results suggest that part of the explanation is that experiments on change blindness almost invariably present observers with unfamiliar scenes. For example, Rensink et al. (2003) used a real world picture of a couple (See Figure 5.1) for the flicker setup and none of their observers saw that exact particular view in their previous experiences. Since observers have no stored model of what these scenes *should* look like, observers are forced to deploy attention serially through the scene and use limited capacity working memory.

For additional support, we can look at the model proposed by Schneider & Shiffrin (1977). These authors described human performance as the result of two qualitatively different forms of information processing, namely automatic and controlled processing. Controlled processing is serial in nature, requires effort, is under an individual's direct control, and requires little or no practice for asymptotic performance. Automatic processing is parallel in nature, not limited by short-term memory capacity, requires little or no effort, is not under a person's direct control and requires extensive consistent training to develop. This distinction also favors our conclusion since the early observation of the change blindness demonstrations include unfamiliar scenes and therefore requires serial / controlled processing (i.e. limited capacity of action). When you become familiar with that scene, on the other hand, you realize the change and it becomes an automatic process since you somehow trained yourself with the image. This explanation would account for missing certain, somewhat arbitrary, changes such as a man's trousers changing color.

Figure 5-1 An Example of a Trial from the "flicker paradigm" (Rensink et al., 2003). In this example, the change is a position change—the horizontal bar moves upward.

Note that a substantial portion of the participants's fixations fell on pedestrians, rather than on the stationary objects, in the present experiments (see Figures 4-10 & 4-11 & 4-22 & 4-23). In previous studies, furthermore, participants were found to be very good at detecting briefly presented Stop signs in a virtual driving environment if the Stop signs were located at intersections. They were much less effective at detecting signs in less likely locations, such as in the middle of a block (Shinoda et al, 2001). Similarly, participants appear to detect potential collisions by actively monitoring other pedestrians in peripheral vision (Jovancevic et al, 2006). Thus the distribution of gaze in a scene is strongly influenced by both the participant's learnt agenda, and deviations from previously learnt state. When this deviation gets larger, it becomes much more difficult to be aware of the scene and its components. Therefore, we may argue that using unfamiliar scenes can make the changes unnoticeable for the observers.

5.2.4 Puzzling Aspects of the Results

There are some slightly puzzling aspects of the results of the first experiment. When objects are removed from the scene, the time spent fixating the region where the object had been was less than the time spent fixating stable objects. This implies that objects that disappear are not prioritized by attention in the absence of retinal transients, consistent with Brockmole and Henderson (2005b) since they claimed that the appearance of an object may be more cognitively salient than the removal of an object. On the other hand, more time was spent fixating these locations when the participant was familiar with the environment. This effect of familiarity suggests that the disappearance of the object does indeed lead to attentional prioritization.

A similar result holds for objects that are replaced, although the difference between the groups is less reliable. It is not clear why the Inexperienced participants looked at the replaced object less than stable objects. It is possible that there is some uncontrolled factor, such as object size, that accounts for the reduced likelihood of a fixation on the replaced object. Note that the billboard and house, two of the stable objects, are much larger than the dog or the fire hydrant, consistent with this speculation. This, and other idiosyncracies of the design limit the conclusions about the specific difference that was observed between the different types of change. For example, the object that moved to a new location was the same object that was novel in the preceding circuit. A fixation on that object when it moved might therefore not be independent of the events on the previous lap.

When the results of the two experiments are compared, we see a conflicting result about the effect of computer usage habits. While the participants' weekly usage of the computers did not seem to influence the performance of the ones in the first experiment, that habit of the participants in the second experiment resulted with a difference in the performance. This may be due to the nature of the virtual environment, since in the second experiment the virtual experience was more similar to a computer game rather than walking in a novel environment as in the first experiment.

5.3 Conclusion

The results of the experiments revealed that participants are more likely to fixate changed objects if they are familiar with the environment. Participants familiar with the environment fixated changes for a significantly longer time period than participants who were less familiar with the environment. These results are consistent with the results of Brockmole and Henderson (2005a), which was conducted with 2D images. The results support the hypothesis that we learn the structure of natural scenes over time, and that attention is attracted by deviations from the normal state.

Results also provided us with the finding that the type of change did not have an interaction with change detection; the average fixation durations on each type of changing object was significantly different from the average fixation durations on its respective baseline for control condition. On the other hand, even if the type of change seems to have no interaction with change detection, it has actually a significant effect on both *Experienced* and the *Inexperienced* type of groups of the second experiment when we look at the average time durations for the first fixation to be allocated on the changed object after it becomes visible to the observer. It means familiarity influences change detection in every type of change even if the nature of these changes are indeed different from each other. This finding also shows the power of learning (i.e. familiarity in our sense) on gaze distribution since it has its effects on each different change type regardless of the nature of change.

Previous results showed that object additions and deletions were fixated at rates greater than chance, suggesting that both types of scene change are cues used by the visual system to guide attention during scene exploration, although appearances were fixated twice as often as disappearances, indicating that new objects are more salient than deleted objects (Brockmole & Henderson, 2005). Other results in the literature also suggested that participants are less likely to detect deletions than to detect additions (Pezdek et al., 1988). In the experiments of this study, I further investigated this issue by adding two different change types; replacement and movement. The results helped us to have more clear judgments about the nature of this distinction as well as the prioritization of other two change types (i.e. replacement & movement). Present results showed that object deletion is also a powerful change type like the other three. Previous underestimation of deletions may due to the control groups that were used for comparisons since empty spots that occurred after deletions were compared with objects in those studies. Comparing the gaze allocations on empty places with that on objects is not cognitively plausible since empty places cannot be as informative as the objects.

The *new-object* condition was found to be the most powerful change type to attract the gaze of the participants while the *deleted-object* condition was the least powerful change type in both experiments. Average fixation duration on the changed object was shorter than the fixation durations on other types of changed objects for the *deleted-object* condition but it must be noted that this is because of the cognitive nature of the process since looking at a present object is much more informative than looking at an empty space. After when the event of disappearance is recognized the viewer may not need to look at that empty spot for gathering further information while this is not the case for other type of changes.

The results of the experiments were also expected to show the differences between the feeling of presence in desktop and immersive virtual environments according to participants' level of awareness for the changes that happen in the environment. When the data are examined in this manner, it can be claimed that no significant difference was observed between the two types of technologies. Furthermore, fixation distribution data indicated a similarity about participants' movement or walking styles between the immersive and desktop virtual reality environments and their common demands with moving in real world places.

The results also have significance for understanding change blindness. The present results, and those of Brockmole and Henderson (2005a), suggest that part of the explanation is that experiments on change blindness almost invariably present observers with unfamiliar scenes. Even if the observers could seen a number of similar scenes before, they cannot have a true idea of what the property or the location of the changed object should be in that specific scene. Since observers have no stored model of what the scenes *should* look like, observers are forced to deploy attention seriallly through the scene and use limited capacity working memory. This explanation would account for missing certain, somewhat arbitrary, changes such as a man's trousers changing color.

5.4 Limitations & Further Study

In order to be able to generalize the results of an experimental study, the number of participants and their demographic diversity should be as large as possible. In this study, the main limitation was the lack of demographic diversity since only university students were participated in the experiments. Although, the number of participants was large enough to see the significant effects, more participants would allow us to see different trends.

Another lack in this study was the impossibility of using the same virtual environment for conducting the experiments. This was due to the demand for using different software programs to create the environments. Since these software programs use different object files it became impossible to use the same objects in the experiments of the study. It could be more reliable to use the same environment for comparing immersive and desktop virtual reality technologies.

Potential track losses due to the movement of the helmet in the first experiment and possible noise of the eye-trackers in both experiments could have possibly affected the results, even if these effects were tried to be minimized. Such effects should be measured and evaluated in a future study.

In a further study, different media or visualization technologies, such as Augmented Reality, CAVE[17], etc., can be used to understand their possible advantages and disadvantages. More change types can be added to the

[17] A Virtual Reality Eenvironment where some or all of the walls of a room are rear-projection stereo displays. The user wears glasses to enable viewing the stereo images, and there is a head tracking mechanism to control what is projected (i.e., the view) depending on where the viewer is located and looking. In addition, there is some mechanism for interacting with what is seen (Buxton & Fitzmaurice, 1998).

procedure for a deeper understanding of different natural tasks. The feeling of presence can also be compared between different experimental settings in order to have more accurate findings about its effects on change detection performance. Furthermore, these results can be compared with real-world cases after the possible future improvements on eye-tracking technology for using it in large scale outdoor environments.

REFERENCES

Anderson, J. R. (1995). Learning and Memory. Toronto: John Wiley & Sons, Inc.

Arthur, H. J. & Hancock, P. A. (2001). Navigation Training in Virtual Environments. *International Journal Of Cognitive Ergonomics,* 5(4), 387–400.

Beck, M. R., Levin, D. T. & Angelone, B. (2007). Change blindness blindness: Beliefs about the roles of intention and scene complexity in change detection. *Consciousness and Cognition*, 16, 31–51.

Berger, A., Henik, A. & Rafal, R. (2005). Competition Between Endogenous and Exogenous Orienting of Visual Attention. *Journal of Experimental Psychology: General* 134(2), 207–221.

Boot, W. R., Kramer, A. F. & Peterson, M. S. (2005). Oculomotor consequences of abrupt object onsets and offsets: Onsets dominate oculomotor capture. *Perception & Psychophysics,* 67(5), 910-928.

Bowman, D. A., Davis, E. T., Hodges, L. F.& Badre, A. N. (1999). Maintaining Spatial Orientation during Travel in an Immersive Virtual Environment. *Presence*, 8(6), 618–631.

Breazeal, C. & Scassellati, B. (1999). A context-dependent attention system for a social robot. In *IJCAI '99: Proceedings of the Sixteenth International Joint Conference on Artificial Intelligence*, 1146-1153, San Francisco, CA, USA.

Brockmole, J. R. & Henderson, J. M. (2005a). Prioritization of New Objects in Real-World Scenes: Evidence from Eye Movements. *Journal of Experimental Psychology / Human Perception & Performance*, 31(5), 857-868.

Brockmole, J. R. & Henderson, J. M. (2005b). Object appearance, disappearance, and attention prioritization in real-world scenes. *Psychonomic Bulletin & Review, 12,* 1061-1067.

Brockmole, J. R. & Henderson, J. M. (2006). Recognition and attention guidance during contextual cueing in real-world scenes: Evidence from eye movements. *Quarterly Journal of Experimental Psychology*, 59(7), 1177-1187.

Buckner R. L. & Wheeler M. E. (2001). The cognitive neuroscience of remembering. *Natural Review of Neuroscience*, 2, 624–634.

Buswell, G. T. (1935). *How people look at pictures. A study of the psychology of perception in art.* The University of Chicago Press, Chicago.

Buxton, W. & Fitzmaurice, G.W.(1998). HMD's, Caves & Chameleon: A Human-Centric Analysis of Interaction in Virtual Space. *Computer Graphics: The SIGGRAPH Quarterly,* 32(4), 64-68.

Cabeza, R. & Nyberg, L. (2000). Imaging cognition. II. An empirical review of 275 PET and fMRI studies. *Journal of Cognitive Neuroscience*, 12(1), 1–47.

Cartwright-Finch, U. & Lavie, N. (2007). The role of perceptual load in inattentional blindness. *Cognition*, 102, 321–340

Casaña Pérez, S. M. (2004). *Control of Attention Before Reflexive And Intentional Saccades.* Doctoral Dissertation, Aus der Neurologischen Klinik und Poliklinik Großhadern der Universität München.

Cech, C. G. (1998). *Principles of Learning* (Ch.1). Retrieved June 13, 2007, from University of Louisiana at Lafayette Web site: http://www.ucs.louisiana.edu/~cgc2646/LRN/Chap1.html.

Chastain, G., Cheal, M. & Kuskova, V. (2002). Inappropriate capture by diversionary dynamic elements. *Visual Cognition, 9(3),* 355-381.

Chen, C., Czerwinski, M. & Macredie, R. (2000) Individual differences in virtual environments – introduction and overview. *Journal of the American Society for Information Science*, 51(6), 499-507.

Chua, H. F., Boland, J. E. & Nisbett, R. E. (2005). Cultural variation in eye movements during scene perception. *Proceedings of the National Academy of Sciences*, USA, 102, 12629-12633.

Chun, M. M. & Jiang, Y. (1998). Contextual cueing: Implicit learning and memory of visual context guides spatial attention. *Cognitive Psychology*, 36, 28-71.

Chun, M. M. & Jiang, Y. (2003). Implicit, long-term spatial contextual memory. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **29**, 224-234.

Cline, D., Hofstetter, H. W. & Griffin, J. R. (1997). *Dictionary of Visual Science*. 4th ed. Boston:Butterworth-Heinemann.

Connor, C. E., Egeth, H. E. & Yantis, S. (2004). Visual Attention: Bottom-Up Versus Top-Down. *Current Biology,* 14, 850-852.

Darken, R.P. (1996). *Wayfinding in Large-Scale Virtual Worlds*. Unpublished doctoral dissertation, The George Washington University, Washington, D.C.

Dennett, D. C. (1991). *Consciousness explained*. London: Penguin Books.

Dikici, C. (2004). *Fovea Based Coding For Video Streaming*. Unpublished MS thesis, Bogazici University, Istanbul.

Droll, J., Hayhoe, M., Triesch, J. & Sullivan, B.(2005). Task demands control acquisition and maintenance of visual information. *Journal of Experimental Psychology: Human Perception and Performance*. 31(6):1416-1438.

Duchowski, A. T., Shivashankaraiah, V., Rawls, T., Gramopadhye, A. K., Melloy, B. J. & Kanki, B. (2000). Binocular eye tracking in virtual reality for inspection training. In A. Duchowski (Ed.): *Proceedings of the International Conference Eye Tracking Research & Applications, ETRA, Palm Beach Gardens, FL, ACM Press,* 89-96.

Duchowski, A. T. (2003). *Eye Tracking Methodology: Theory and Practice*. (3-15) London: Springer.

Faubert, J. (2001) Motion parallax, stereoscopy, and the perception of depth: Practical and theoretical issues. *Proceedings of SPIE,* CR76, 168-191.

Fernandez-Duque, D. & Thornton, I. M. (2000). Change Detection Without Awareness: Do Explicit Reports Underestimate the Representation of Change in the Visual System? *Visual Cognition,* 7, 323-344.

Folk, C.L., Remington, R.W. & Johnston, J.C., (1992). Involuntary covert orienting is contingent on attentional control settings. *Journal of Experimental Psychology: Human Perception and Performance.* 18(4), 1030–1044.

Franconeri, S. L. & Simons, D. J. (2003). Moving and looming stimuli capture attention. *Perception & Psychophysics*, 65(7), 999-1010.

Gabrieli, J. D. E. (1998). Cognitive neuroscience of human memory. *Annual Review of Psychology*, 49**,** 87–115.

Gattis, M. (2001). *Spatial Schemas and Abstract Thought* (1-12). Cambridge, MA: MIT Press.

Gibson, B. S. & Jiang, Y. (1998). Surprise! An unexpected color singleton does not capture attention in visual search. *Psychological Science*, 9(3), 176-182.

Gibson, B.S. & Kelsey, E.M. (1998). Stimulus-driven attentional capture is contingent on attentional set for displaywide visual features. *Journal of Experimental Psychology: Human Perception and Performance*, *24(3)*, 699–706.

Goujon, A., Didierjean, A. & Marméche, E. (2007). Contextual cueing based on specific and categorical properties of the environment. *Visual Cognition*, 15(3), 257-275.

Gregory, R. L. (1973). *Eye and brain; the psychology of seeing* (42-44). New York: McGraw-Hill.

Grimes, J. (1996) On the failure to detect changes in scenes across saccades. In K. Akins (Ed.), *Perception* (89-110). Oxford University Press.

Gur, M. & Snodderly, D.M. (1997). Visual Receptive Fields of Neurons in Primary Visual Cortex (VI) Move in Space with the Eye Movements of Fixation. *Vision Research,* 37(3), 257–265.

Haber, R. N. & Hershenson, M. (1973). *The Psychology of Visual Perception*. (112-114 & 247-273). New York: Holt, Rinehart and Winston, Inc.

Hayhoe, M. M., Bensinger, D.G. & Ballard, D.H. (1998). Task constraints in visual working memory. *Vision Research*, *38(1)*, 125–137.

Hayhoe, M. M., Shrivastava, A., Mruczek, R. & Pelz, J.B. (2003). Visual memory and motor planning in a natural task. *Journal of Vision,* 3(1), 49-63.

Hayhoe, M. M. & Ballard, D. (2005). Eye movements in natural behavior. *Trends in Cognitive Sciences, 9*(4), 188-193.

Henderson, J. M., & Hollingworth, A. (1999). The role of fixation position in detecting scene changes across saccades. *Psychological Science, 10,* 438-443.

Hollingworth, A. & Henderson, J. M. (2000). Semantic Informativeness Mediates the Detection of Changes in Natural Scenes. *Visual Cognition*, *7* (1/2/3), 213–235.

Hollingworth, A., Schrock, G. & Henderson, J. M. (2001). Change detection in the flicker paradigm: The role of fixation position within the scene. *Memory & Cognition,* 29 (2), 296-304.

Hollingworth, A. & Henderson, J. M. (2002). Accurate visual memory for previously attended objects in natural scenes. *Journal of Experimental Psychology: Human Perception and Performance, 28,* 113-136.

Hollingworth, A. & Henderson, J. M. (2004). Sustained change blindness to incremental scene rotation: A dissociation between explicit change detection and visual memory. *Perception & Psychophysics, 66,* 800-807.

Hyrskykari, A., Majaranta, P., Aaltonen, A. & Raiha, K. (2000). Design Issues of iDict: A Gaze-Assisted Translation Aid. In A. Duchowski (Ed.): *Proceedings of the International Conference Eye Tracking Research & Applications, ETRA, Palm Beach Gardens, FL, ACM Press,* 9-15.

Irwin, D. E. & Andrews, R. (1996). Integration and accumulation of information across saccadic eye movements. In T. Inui & J. L. McClelland (Eds.), *Attention and performance: XVI. Information integration in perception and communication* (125-155). Cambridge, MA: MIT Press.

Irwin D.E., Colcombe A.M., Kramer A.F. & Hahn S. (2000). Attentional and oculomotor capture by onset, luminance and color singletons. *Vision Research*, 40(10-12), 1443–1458.

Itti, L. (2000) *Models of bottom-up and top-down visual attention*. Unpublished doctoral dissertation, California Institute of Technology, California.

Itti, L. & Koch, C. (2001) Computational Modeling of Visual Attention. *Nature Reviews Neuroscience* 2(3), 194-203.

Itti, L., Gold, C. & Koch, C. (2001) Visual Attention and Target Detection in Cluttered Natural Scenes, *Optical Engineering,* 40 (9), 1784-1793.

Itti, L., Koch, C. & Niebur, E.(1998). A model of saliency-based visual attention for rapid scene analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(11), 1254-1259.

Jacoby, L.L. (1992). Strategic versus automatic influences of memory-attention, awareness and control. *Bulletin of the Psychonomic Society*, *30*, 442-443.

Jiang, Y. & Song, J-H (2005). Spatial context learning in visual search and change detection. *Perception & Psychophysics, 67(7), 1128-1139.*

Jiang, Y., Song, J.-H. & Rigas, A. (2005). High-capacity spatial contextual memory. *Psychonomic Bulletin & Review*, 12, 524-529.

Jonides, J. & Yantis, S. (1988). Uniqueness of abrupt visual onset in capturing attention. *Perception & Psychophysics*, *43*, 346–354.

Jovancevic, J., Hayhoe, M. & Sullivan, B. (2004). Control of gaze while walking [Abstract]. *Journal of Vision*, *4*(8), 4a.

Jovancevic, J., Sullivan, B. & Hayhoe, M. (2005). Attentional capture for potential collisions gated by task [Abstract]. *Journal of Vision*, *5*(8), 101a.

Jovancevic, J., Sullivan, B. & Hayhoe, M. (2006). Learning gaze allocation priorities in complex environments [Abstract]. *Journal of Vision*, *6*(6), 480a.

Kamas, E. & Reder, L.M. (1994).The role of familiarity in cognitive processing. In: E. O'Brien, and R. Lorch (Eds.), *Sources of coherence in reading: A festschrift in honor of Jerome L. Myers.* New Jersey: L. Erlbaum, 177-202.

Karacan (Üke), H. & Hayhoe, M. M. (accepted) Is attention drawn to changes in familiar scenes? *Visual Cognition*.

Keat, J., Balendran, V., Sivayoganathan, K. & Sackfield, T. (1999). An Adaptive Connectionist Front-End for Automated Fettling. *Integrated Computer-Aided Engineering*, 6(1), 67-89.

Koch, C. & Ullman, S. (1985). Shifts in selective visual attention: towards the underlying neural circuitry. *Human Neurobiology*, 4, 219-227.

Krieger, G., Rentschler, I., Hauske, G., Schill, K. & Zetzsche, C. (2000). Object and scene analysis by saccadic eye-movements: an investigation with higher-order statistics. *Spatial Vision* 13, 201-214.

Khadhouri, B. & Demiris, Y. (2005) Compound effect of bottom-up and top-down influences in attention during action recognition*, Proceedings of IJCAI, Edinburgh*.

Koesling, H., Carbone, E. & Turku, H. R. (2001). VDesigner – A Visual Programming Environment for Eye Tracking Experiments. *Proceedings 11th European Conference on Eye Movements, ECEM11, Finland*.

Land, M. F. & Lee, D. N. (1994). Where we look when we steer. *Nature,* 369, 742-744.

Land, M. & Furneaux, S. (1997). The knowledge base of the oculomotor system. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 352, 1231–1239.

Land, M., Mennie, N. & Rusted, J. (1999). The roles of vision and eye movements in the control of activities of daily living. *Perception,* 28:1311--1328.

Land, M. F. & McLeod, P. (2000). From eye movements to actions: How batsmen hit the ball. *Nature Neuroscience,* 3, 1340-1345.

Land, M. F. & Hayhoe, M. M. (2001). In what ways do eye movements contribute to everyday activities. *Vision Research*, 41 (25-26), 3559-3565.

Le Meur, O., Le Callet, P., Barba, D. & Thoreau, D. (2006) A Coherent Computational Approach to Model Bottom-Up Visual Attention. *IEEE Transactions on Pattern Analysis and Machine Intelligence,* 28 (5), 802-817.

Longman Dictionary of contemporary English (1995). New edition. Harlow: Longman.

Loomis, J., Blascovich, J. & Beall, A. (1999). Immersive virtual environment Technology as a basic research tool in psychology. *Behavior Research Methods, Instruments, & Computers* , 31 (4), 557 - 564.

Mack, A., & Rock, I. (1998). *Inattentional Blindness, Ch.1.* Cambridge, MA: MIT Press.

Mandler, G. (1980). Recognising: The judgement of a previous occurrence. *Psychological Review*, *27*, 252-271.

Mannan S.K, Ruddock K.H. & Wooding D.S. (1997). Fixation sequences made during visual examination of briefly presented 2D images. *Spatial Vision* 11, 157- 178.

Marr, D. (1982). *Vision: A computational investigation into the human representation and processing of visual information*. (8-41). San Francisco: Freeman.

May, J. G. & Badcock, D. R. (2002). Vision and Virtual Environments. In Stanney, K. M. (Eds.) *Handbook of Virtual Environments: Design, Implementation, and Applications.* (29-65). New Jersey: Lawrence Erlbaum Associates.

Mazuryk, T. & Gervautz, M. (1996). Virtual reality. history, applications, technology and future, Research Paper, Institute of Computer Graphics, Vienna University of Technology. Retrieved Online October 17, 2006, from: http://www.cg.tuwien.ac.at/research/publications/1996/mazuryk-1996-VRH/TR-186-2-96-06Paper.pdf

McConkie, G.W. & Rayner, K. (1975). The span of the effective stimulus during a fixation in reading, *Perception & Psychophysics* 17, 578–586.

Melcher, D. & Kowler, E. (2001). Visual scene memory and the guidance of saccadic eye movements. *Vision Research*, 41(25-26), 3597-3611.

Melcher, D. (2005). Spatiotopic transfer of visual-form adaptation across saccadic eye movements. *Current Biology*, 15(19), 1745-1748.

Milgram, P., Takemura, H., Utsumi, A. & Kishino, F. (1994). Augmented Reality: A class of displays on the reality-virtuality continuum. *SPIE Vol. 2351-34, Telemanipulator and Telepresence Technologies*.

Moffat, S.D., Zonderman, A.B. & Resnick, S.M. (2001). Effects of age on spatial learning in a virtual reality navigation task. *Neurobiology of Aging,* 22, 787-796.

Mondy, S. & Coltheart, V. (2000). Detection and Identification of Change in Naturalstic Scenes. *Visual Cognition*, 7(1/2/3), 281-296.

Monin, B. (2003). The Warm Glow Heuristic: When Liking Leads to Familiarity. *Journal of Personality and Social Psychology*, 85(6), 1035–1048.

Montaldi, D., Spencer, T.J., Roberts, N. & Mayes, A.R. (2006). The Neural System That Mediates Familiarity Memory. *Hippocampus*, 16, 504–520.

Moran, A. (2000). Exploring visual attention using eye-tracking technology. Retrieved October 04, 2006, from The Irish Scientist 2000 Year Book Web site: http://www.irishscientist.ie/2000/contents.asp?contentxml=203bs.xml&contentxsl=insight3.xsl

Munro, A., Breaux, R., Patrey, J. & Shelton, B. (2002). Cognitive Aspects of Virtual Environments Design. In Stanney, K. M. (Eds.) *Handbook of Virtual Environments: Design, Implementation, and Applications.* (415-435). New Jersey: Lawrence Erlbaum Associates.

Murray E. A.& Richmond, B. J. (2001). Role of perirhinal cortex in object perception, memory, and associations. *Current Opinion in Neurobiology*, 11(2)**,** 188–193.

Naish, P.L.N. (2005). Attention. In Braisby, N., Gellatly, A. (Eds.), *Cognitive Psychology* (34-67). New York, NY: Oxford University Press.

Noe, A., Pessoa, L. & Thompson, E. (2000). Beyond the Grand Illusion: What Change Blindness Really Teaches Us About Vision. *Visual Cognition*, *7* (1/2/3), 93–106.

Oliva, A., Torralba, A., Castelhano, M. & Henderson, J. (2003). Top Down Control of Visual Attention in Object Detection. *Journal of Vision*, 3(9), 3a.

Onay-Durdu, P., Çağıltay, K. (2006). Investigation of Simulator Sickness in A Driving Simulator. Proceedings of *Traffic and Road Safety Third International Congress*, Gazi University, Ankara.

O'Regan J.K., Rensink R.A., & Clark J.J. (1999). Change blindness as a result of 'mudsplashes'. *Nature,* 398, 34.

Palmer, J. (1990). Attentional Limits on the Perception and Memory of Visual Information. *Journal of Experimental Psychology: Human Perception and Performance*, 16(2), 332-350.

Parkust, D. J., Culurciello, E. & Niebur, E. (2000). Evaluating Variable Resolution Displays with Visual Search: Task Performance and Eye Movements. In A. Duchowski (Ed.): *Proceedings of the International Conference Eye Tracking Research & Applications, ETRA, Palm Beach Gardens, FL, ACM Press,* 105-111.

Parkhurst, D. J. & Niebur, E. (2003). Scene content selected by active vision. *Spatial Vision,* 16(2), 125-154.

Patla, A. E. & Vickers, J. N. (1997). Where and when do we look as we approach and step over an obstacle in the travel path? *Neuroreport: An International Journal for the Rapid Communication of Research in Neuroscience, 8*, 3661-3665.

Pelz, J. B., Canosa, R. & Babcock, J. (2000). Extended Tasks Elicit Complex Eye Movement Patterns. In A. Duchowski (Ed.): *Proceedings of the International Conference Eye Tracking Research & Applications, ETRA, Palm Beach Gardens, FL, ACM Press,* 37-45.

Pelz, J.B. & Canosa, R. (2001). Oculomotor behavior and perceptual strategies in complex tasks. *Vision Research*, 41(25-26), 3587–3596.

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

Pike, G. & Brace, N. (2005). Recognition. In Braisby, N., Gellatly, A. (Eds.), *Cognitive Psychology* (113-153). New York, NY: Oxford University Press.

Pike, G. & Edgar, G. (2005). Perception. In Braisby, N., Gellatly, A. (Eds.), *Cognitive Psychology* (71-109). New York, NY: Oxford University Press.

Posner, M. I. (2003). Attention: The Mechanisms of Consciousness. In B.J. Baars, W.P. Banks & J.B. Newman (Ed.): *Essential Sources in the Scientific Study of Consciousness* (279-289). Cambridge: MIT Press.

Rao, R. P. N. & Ballard, D. H. (1998). Development of localized oriented receptive fields by learning a translation-invariant code for natural images. *Network: Computation in Neural Systems*, 9(2), 219-234.

Rao, R. P., Zelinsky, G., Hayhoe, M. & Ballard, D. H. (2002). Eye movements in iconic visual search. *Vision Research,* 42(11),1447-1463.

Reinagel, P., Godwin, D., Sherman, S. M. & Koch, C. (1999). Encoding of visual information by LGN bursts. *Journal of Neurophysiology,* 81, 2558–2569.

Rensink, R. A., O'Regan, J. K. & Clark, J. (2003). To see or not to see: the need for attention to perceive changes in scenes. In B.J. Baars, W.P. Banks & J.B. Newman (Ed.): *Essential Sources in the Scientific Study of Consciousness* (251-261). Cambridge: MIT Press.

Rensink, R. A. (2000). The dynamic representation of scenes. *Visual Cognition, 7,* 17-42.

Rensink, R. A. (2002). Change detection. *Annual Review of Psychology*, 53, 245–277.

Richardson, A. E., Montello, D. & Hegarty, M. (1999). Spatial knowledge acquisition from maps, and from navigation in real and virtual environments. *Memory & Cognition*, *27*, 741-750.

Richardson, D.C., Dale, R. & Spivey, M.J. (2007). Eye Movements in Language and Cognition: A Brief Introduction. In M. Gonzalez-Marquez, I. Mittelberg, S. Coulson & M. J. Spivey (Ed.): *Empirical Methods in Cognitive Linguistics* (323-345). Amsterdam: John Benjamin.

Rothkopf, C., Ballard, D, Sullivan, B, & de Barbaro, K. (2005) Bayesian modeling of task dependent visual attention strategy in a virtual reality environment [Abstract]. *Journal of Vision*, *5(8)*,920a.

Rutishauser, U., Walther, D., Koch, C. & Perona, P. (2004). *Is bottom-up attention useful for object recognition*? *Proceedings of the 2004 IEEE International Conference on Computer Vision and Pattern Recognition (CVPR),* (2), 37-44.

Schmuckler, M. A.(2001). What Is Ecological Validity? A Dimensional Analysis. *Infancy*, 2(4), 419-436.

Schneider, W. & Shiffrin, R.M. (1977). Controlled and automatic human information processing: I. Detection, search, and attention. *Psychological Review, 84*(1), 1-66.

Scholl, B. J. (2000). Attenuated change blindness for exogenously attended items in a flicker paradigm. *Visual Cognition, 7,* 377-396.

Sherman, W.R. & Craig, A.B. (2003). *Understanding Virtual Reality: Interface, Application and Design.* (4-36). San Francisco, CA : Morgan Kaufmann.

Shinoda, H., Hayhoe, M. & Shrivastava, A. (2001). What controls attention in natural environments. *Vision Research,* 41(25-26), 3535-3545.

Simons, D. J. & Levin, D. T. (1997). Change Blindness. *Trends in Cognitive Sciences, 1*(7), 261-267.

Simons, D. J. (2000a). Current approaches to change blindness. *Visual Cognition: Special Issue on Change Detection & Visual Memory*, 7, 1-16.

Simons, D. J. (2000b). Attentional Capture and Inattentional Blindness. *Trends in Cognitive Sciences*, 4(4), 147-155.

Simons, D. J. & Ambinder, M. S. (2005). Change Blindness. *Current Directions in Psychological Science* 14 (1), 44-48.

Simons, D. J. & Rensink, R. A. (2005). Change Blindness: Past, Present, and Future. *Trends in Cognitive Sciences*, 9 (1), 16-20.

Stankiewicz, B.J., Legge, G.E., Mansfield, J.S. & Schlicht, E.J. (2006). Lost in Virtual Space: Studies in Human and Ideal Spatial Navigation. *Journal of Experimental Psychology: Human Perception & Performance.* 32(3), 688–704.

Stein, F. J. (1992). *Structural Indexing for Object Recognition*. Unpublished doctoral dissertation, University of Southern California, California.

Szara, S.I. & Lutford, J.P. (1981). Benzodiazepines: A Review of Research Results, 1980. NIDA Research Monograph 33, A RAUS Review Report.

Takemura, H.& Kishino, F. (1992). Cooperative work environment using virtual workspace. *Proceedings of Computer Supported Cooperative Work (CSCW'92)*, 226-232.

Theeuwes J. (1991). Exogenous and endogenous control of attention: The effect of visual onsets and offsets. *Perception & Psychophysics*, 49(1), 83-90.

Theeuwes, J. (1994). Stimulus-driven capture and attentional set: Selective search for color and visual abrupt onsets. *Journal of Experimental Psychology: Human Perception & Performance, 20,* 799-806.

Theeuwes J., Kramer A.F., Hahn S. & Irwin D.E. (1998). Our eyes do not always go where we want them to go: capture of the eyes by new objects. *Psychological Science*, 9(5), 379–385.

Theeuwes J. (2004). Top-down search strategies cannot override attentional capture. *Psychonomic Bulletin & Review,* 11 (1), 65-70.

Torralba, A. (2003). Contextual priming for object detection *International Journal of Computer Vision,* 53 (2), 169–91.

Treisman, A. & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology,* 12 (1), 97-136.

Treisman, A. (2006). How the deployment of attention determines what we see. *Visual Cognition*, 14 (4-8), 411-443.

Triesch, J., Sullivan, B. T., Hayhoe, M. M., & Ballard, D. H. (2002). Transient visual representations: a change blindness approach [Abstract]. *Journal of Vision*, *2*(7), 244a.

Triesch, J., Ballard, D., Hayhoe, M. & Sullivan, B. (2003). What you see is what you need. *Journal of Vision,* 3, 86-94.

Tsotsos, J. K. (2001). Motion understanding: Taskdirected attention and representations that link perception with action. *International Journal of Computer Vision*, *45 (3), 265-280*.

Turano, K. A., Geruschat, D. R. & Baker, F. H. (2003). Oculomotor strategies for the direction of gaze tested with a real-world activity. *Vision Research, 43(3)*, 333-346.

Turatto, M. & Galfano, G. (2000). Color, form and luminance capture attention in visual search. *Vision Research*, *40(13)*, 1639–1643.

Underwood, G. & Bright, J. E. H. (1996). Cognition with and without awareness. In G. Underwood (Ed.): *Implicit Cognition*, New York: Oxford University Press, 1-41.

Üke, H. (2005). *The Role of Attention for Visual Perception in Desktop Virtual Reality.* Unpublished master thesis, Middle East Technical University, Ankara.

Üke, H. & Hayhoe, M. (2006). Is attention drawn to changes in familiar scenes? [Abstract]. *Journal of Vision*, *6*(6), 1095a.

Velichkovsky, B. M., Dornhoefer, S. M., Pannasch, S. & Unema, P.J.A. (2001). Visual fixations and level of attentional processing. In A. Duchowski (Ed.): *Proceedings of the International Conference Eye Tracking Research & Applications, ETRA, Palm Beach Gardens, FL, ACM Press,* 79-87.

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.

Waller, D. (2000). Individual differences in spatial learning from computer-simulated environments. *Journal of Experimental Psychology: Applied*, 8, 307 – 321.

Wang, R. F. & Brockmole, J. R. (2003). Simultaneous spatial updating in nested environments. *Psychonomic Bulletin & Review*, 10, 981-98.

Welch, G., Bishop, G., Vicci, L., Brumback, S., Keller, K., & Colucci, D. (1999). The HiBall tracker: High-performance wide-area tracking for virtual and augmented environments. In *Symposium on Virtual Reality Software and Technology*, 1-10.

Wilder, J., Hung, G. K., Tremaine, M. & Kaur, M. (2002). Eye-Tracking in Virtual Environments. In Stanney, K. M. (Eds.) *Handbook of Virtual Environments: Design, Implementation, and Applications* (211-222). New Jersey: Lawrence Erlbaum Associates.

Witmer, B. G., Bailey, J. H., Knerr, B. W. & Parsons, K. C. (1996). Virtual spaces and real world places: transfer of route knowledge. *International Journal of Human-Computer Studies,* 45(4), 413-428.

Wolfe, J. M. (1999). Inattentional Amnesia. In V. Coltheart (Ed.), *Fleeting Memories* (71-94). Cambridge: MIT Press.

Wright, M.J. (2005). Saliency predicts change detection in pictures of natural scenes. *Spatial Vision, 18(4)*, 413-430.

Vince, J. (1999). *Essential virtual reality fast: how to understand the techniques and potential of virtual reality. (27-91).* Great Britain: Springer.

Yantis, S. & Jonides J. (1984). Abrupt visual onsets and selective attention: Evidence from visual search. *Journal of Experimental Psychology: Human Perception and Performance*, 10, 601--621.

Yantis, S. (1993). Stimulus-driven attention capture. *Current Directions in Psychological Science*, 2, 156-161.

Yantis, S. & Hillstrom, A.P. (1994). Stimulus-driven attentional capture: Evidence from equiluminant visual objects. *Journal of Experimental Psychology: Human Perception and Performance*, *20*, 95–107.

Yantis, S. & Jonides, J. (1996). Attentional capture by abrupt onsets: New perceptual objects or visual masking? *Journal of Experimental Psychology: Human Perception and Performance*, 22(6), 1505-1513.

Yantis, S. (1998). Control of visual attention. In H. Pashler (Ed.), *Attention*. East Sussex, England: Psychology Press, 233-256.

Yonelinas, A.P. (2002). The Nature of Recollection and Familiarity: A Review of 30 Years of Research. *Journal of Memory and Language* 46**,** 441–517.

Yoo, D.H., Kim, J.H., Lee, B.R. & Chung, M.J. (2002). Non-contact Eye Gaze Tracking System by Mapping of Corneal Reflections. *Proceedings of the Fifth IEEE International Conference on Automatic Face and Gesture Recognition (FGR.02),* 101-106.

Yu, X.S. & Lee, T.S. (2000). What do V1 neurons tell us about saccadic suppression? *Neurocomputing* 32-33, 271-277.

APPENDICES

APPENDIX A – CONSENT FORM FOR EXPERIMENT 1

Project Title: Vision in Natural tasks
RSRB: 2793
Principal Investigator: Mary Hayhoe

Introduction:

This consent form describes a research study and what you may expect if you decide to participate. You are encouraged to read this consent form carefully and to ask the person who presents it any further questions you might have before making your decision whether or not to participate.

You are being asked to participate in a research study because you indicated an interest in doing so, and have normal or corrected to normal visual activity, which makes it easier for us to monitor eye position effectively.

This form describes the known possible risks and benefits. You are completely free to choose whether or not to participate in the study.

Purpose of Study:

You will be participating in an experiment by the principal investigator concerning how people use their hands and eyes when performing simple everyday tasks. The general goals of the project are to understand the normal functioning of human visually guided behavior.

Description of Study Procedures:

Walk around the room in a virtual environment, wearing the helmet.

For some of the experiments, a more detailed description of the particular experiment, and instructions for that task, are provided in a separate form that should accompany this consent form. Each experimental session lasts from 1 to 2 hours. If you wish to participate in any of the other experiments at another time you are free to do so. Please feel free to ask any questions about the experiment.

Eye, head and hand tracking:

In all experiments we will track eye movements. To do this we use a lightweight camera system, worn on the head, or mounted inside the virtual reality helmet. This device uses a video camera using an infrared light mounted near the eye(s) to measure direction of gaze.

In experiment, you will be asked to walk a short distance (about 30 feet) in the experimental room, wearing the virtual reality helmet. In this experiment your head movements will be tracked by an optical motion tracking system. This system consists of a miniature optical device, mounted on the virtual reality helmet that measures its position with respect to any array of small infra-red lights on the ceiling.

Risks of Participation:

The eye tracking devices used in these experiments have been used for many years and have no known adverse effect. The head band of the tracker occasionally causes a mild headache which ceases as soon as the head band is removed. Also while the eye tracking head band setup and virtual reality helmet are relatively light weight some subjects still may feel some discomfort from the added weight on their head. Please tell the experimenter if you are uncomfortable in any way.

There are no known adverse effects of the virtual reality system. Some individuals are sensitive to small visual discrepancies and may feel slight motion sickness. Please tell the experimenter if you feel nauseous or uncomfortable at any time to halt the experiment.

When walking with the helmet on, an experimenter will walk beside you to make sure you do not bump into anything.

Benefits of Participation:

There are no direct benefits to participating in this study.

Voluntary Participation:

Participation in the study is completely voluntary, and you may terminate participation at any time if you wish to do so and still receive compensation. If you withdraw from the study, information you have provided will be kept confidential.

New Findings:

You will be informed of any new findings which may affect your decision to continue to participate in the study.

Circumstances for Leaving the Study:

You may be asked to leave the study if we are unable to track your eye position reliably. This is a common experience in eye tracking experiments and has no implications concerning your vision. In this event you will still be compensated for your time even if we can not complete the experiment.

Payment:

A small hourly payment of $10 per hour will be given for your participation.

Confidentiality of Records:

Subjects' data is usually published in journal articles and subjects are referred to buy their initials. If you do not wish to be identified by you true initials you may provide a set of initials of your choice to the experimenter for use in publications. Only the investigators will have the access to the subject's identity. While we make every effort to maintain your confidentiality it can not be absolutely guaranteed. (Please note, however, that the data we collect concerns natural, everyday movements of the eye, and this is not normally of a sensitive nature.)

Contact Persons:

For more information concerning this research, you should contact the principal investigator at (585) 2758673 or the Chairman of the Department of Brain and Cognitive Sciences at (585) 2751844. If you have any questions about your rights as a research subject, you may contact:
Human Subjects Protection Specialist
University of Rochester Research Subjects Review Board
Box 315, 601 Elmwood Ave, Rochester, NY 14642
Telephone (585) 2760005

Signatures/Dates:

My signature indicates that I have read and understand the above, questions have been satisfactorily answered, and that I consent to participate in the study. I understand that I may withdraw from the experiment at any time, and that I will receive a copy of this form for my records.

Study Subject Name (Please Print): …………………………………………….
Subject's Signature: …………………………….. Date: …………………….
Investigator's signature: ………………………... Date: ………..…………..
Initials: …………………

APPENDIX B – SUBJECT QUESTIONNAIRE FOR EXPERIMENT 1

Please answer the following OPTIONAL demographic questions by marking the box or filling in the blank with the appropriate response

Biographical Information

1. Name: ………………………………
2. Age: ………………………………..
3. Sex:
   - ☐ Male
   - ☐ Female
4. Race:
   - ☐ Black/African American
   - ☐ Native American/Alaskan Native
   - ☐ Hawaiian/Pacific Islander
   - ☐ Hispanic
   - ☐ Asian
   - ☐ White/Caucasian
   - ☐ Other/More than one race: ………………………..

Vision/Motor Information

1. Do you wear glasses/contacts?
   - ☐ Yes-Power (if known): …………………………….
   - ☐ No

2. Do you have any known vision problems? (i.e. Color blindness, stigmatism)

☐ Yes-Explain: …………………

☐ No

3. Which of your hand is dominant?

☐ Left

☐ Right

APPENDIX C – CONSENT FORM FOR EXPERIMENT 2

Bu çalışma katılımcıların üç boyutlu dinamik bir ortamda incelemeler yaparken ortaya çıkan göz tepkilerini analiz etmeye yönelik bir çalışmadır. Çalışmanın amacı göz hareketi bilgilerini toplayarak verilen tepkilerinin genellenebilirliğini ortaya çıkarmaktır. Çalışma yaklaşık olarak 20 dakika sürecek olup katılım tamamiyle gönüllülük temelindedir. Cevaplarınız tamamiyle gizli tutulacak ve sadece araştırmacı tarafından değerlendirilecektir; elde edilecek bilgiler bilimsel yayımlarda kullanılacaktır.

Çalışma genel olarak kişisel rahatsızlık verecek bir yapı içermemektedir. Ancak, katılım sırasında herhangi bir nedenden ötürü kendinizi rahatsız hissederseniz çalışmayı yarıda bırakıp çıkmakta serbestsiniz. Böyle bir durumda uygulayıcı kişiye, devam etmek istemediğinizi söylemek yeterli olacaktır. Bilgi toplama sonunda, bu çalışmayla ilgili sorularınız cevaplanacaktır. Bu çalışmaya katıldığınız için şimdiden teşekkür ederim.

Çalışma hakkında daha fazla bilgi almak için Araş. Gör. Hacer Karacan (210 3747; E-posta: hacer@ii.metu.edu.tr) ile iletişim kurabilirsiniz.

✂- -

Bu çalışmaya tamamen gönüllü olarak katılıyorum ve istediğim zaman yarıda kesip çıkabileceğimi biliyorum. Verdiğim bilgilerin bilimsel amaçlı yayımlarda kullanılmasını kabul ediyorum. (Formu doldurup imzaladıktan sonra uygulayıcıya geri veriniz).

Adı Soyadı	Öğrenci Numarası	Tarih	İmza
____________	________________	___/___/_______	

APPENDIX D – SUBJECT QUESTIONNAIRE FOR EXPERIMENT 2

Yaşınız:

Cinsiyetiniz: ☐ Bayan ☐ Erkek

Bölümünüz:

Kaçıncı sınıftasınız?

☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ Ms ☐ PhD

Ne kadar zamandır bilgisayar kullanıyorsunuz?

☐ 2 yıldan az

☐ 2-4 yıl

☐ 4-6 yıl

☐ 6-8 yıl

☐ 8-10 yıl

☐ 10 yıldan fazla

Haftada kaç saat bilgisayar kullanıyorsunuz?

☐ 10 saatten az

☐ 10-15 saat

☐ 15-20 saat

☐ 20-25 saat

☐ 25-30 saat

☐ 30 saatten fazla

Bilgisayar oyunları oynar mısınız?

☐ Evet ☐ Hayır

Herhangi bir görme probleminiz var mı, varsa nedir?

☐ Evet ________________________

☐ Hayır

APPENDIX E –DEBRIEFING SHEET FOR EXPERIMENT 2

Bu çalışma daha önce de belirtildiği gibi Bilişsel Bilimler Bölümü bünyesinde yürütülmekte olan katılımcıların göz hareketlerinin analizi sonucu dikkatlerinin öğrenmeyle olan olası ilişkisini anlamaya yönelik bir çalışmadır. Çalışma kapsamında toplanan bilgiler video analiz metotları kullanılarak incelenecek ve insanların dikkatlerini çevrelerindeki bir değişikliğe yönlendirmelerinde öğrenmenin bir rolü olup olmadığı ve bunun farklı değişiklik türleri için belli yapılar içerip içermediği araştırılacaktır.

Bu çalışmadan elde edilecek verilerle yapılacak analizlerin Ağustos 2007 sonunda sonuçlandırılması amaçlanmaktadır. Elde edilen bilgiler sadece bilimsel araştırma ve yazılarda kullanılacaktır. Çalışmanın sonuçlarını öğrenmek ya da bu araştırma hakkında daha fazla bilgi almak için aşağıdaki isimlere başvurabilirsiniz. Bu araştırmaya katıldığınız için tekrar çok teşekkür ederim.

Araş. Gör. Hacer KARACAN (210 3747; E-posta: hacer@ii.metu.edu.tr)

CURRICULUM VITAE

PERSONAL INFORMATION

Surname, Name: Karacan (Üke), Hacer

Nationality: Turkish (TC)

Date and Place of Birth: 15 July 1980, Erzurum

Marital Status: Married

Phone: +90 312 210 37 47

Fax: +90 312 210 37 45

email: hacer@ii.metu.edu.tr

EDUCATION

Degree	Institution	Year of Graduation
MS	METU Cognitive Science	2005
BS	METU Computer Education & Instructional Technology	2002

WORK EXPERIENCE

Year	Place	Enrollment
2002- Present	METU Department of Cognitive Science	Research Assistant
2001-2002	METU Faculty of Education	Student Assistant

PUBLICATIONS

1. **Üke, H.** (2005). The Role of Attention in Visual Perception for Desktop Virtual Reality Environments. *Unpublished Master's Thesis, Middle East Technical University.*

2. **Üke, H.** & Hayhoe, M. (2006). Is attention drawn to familiar scenes? [Abstract]. *Journal of Vision*, 6(6), 1088a.

3. **Üke, H.** (2006). The Usability Study of the Trafikent Driver Training Simulator. *Proceedings of Traffic and Road Safety Third International Congress*, 545-553.

4. Baştanlar, Y., Cantürk, D. & **Karacan, H.** (2007). Effects of Color-Multiplex Stereoscopic View on Memory and Navigation. *Poster session presented at 3DTV 2007 Conference, Greece.*

5. **Karacan (Üke), H.** & Hayhoe, M. (accepted). Is attention drawn to familiar scenes? *Visual Cognition*.

RESEARCH INTERESTS

Virtual Reality, Human-Computer Interaction, Visual Perception, Spatial Awareness