UNIVERSAL ACCESS TO INFORMATION TECHNOLOGY FOR OLDER ADULTS WITH VISUAL IMPAIRMENTS

A Thesis
Presented to
the Academic Faculty

by

Virginia Kathlene Leonard

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in Industrial and Systems Engineering

Georgia Institute of Technology
August, 2005

UNIVERSAL ACCESS TO INFORMATION TECHNOLOGY FOR OLDER ADULTS WITH VISUAL IMPAIRMENTS

Approved by:

Dr. Julie A. Jacko, Advisor
School of Industrial & Systems Engineering
Georgia Institute of Technology

Dr. François Sainfort, Co-Advisor
School of Industrial & Systems Engineering
Georgia Institute of Technology

Dr. Brani Vidakovic
School of Industrial & Systems Engineering
Georgia Institute of Technology

Dr. Joseph J. Pizzimenti
College of Optometry
Nova Southeastern University

Dr. Marie Thursby
College of Management
Georgia Institute of Technology

Date Approved: June 17, 2005

For my parents

Richard & Carol Emery

Your unconditional love, support and encouragement

inspire me to succeed; I am forever grateful for you both

ACKNOWLEDGMENTS

First and foremost, I would like to thank my advisor, Dr. Julie Jacko and co-advisor, Dr. François Sainfort. I have grown considerably as an academic and also personally since beginning my work at Georgia Tech, owing in large part to the model of excellence demonstrated by my mentors and the confidence they have always projected in my ability to succeed.

I would like to thank my committee members, Dr. Brani Vidakovic, Dr. Marie Thursby and Dr. Joseph Pizzimenti for their collaborations, invaluable feedback and enthusiasm for this project during both the dissertation process and over the course of my academic career at Georgia Tech.

This research would not have been possible without the support of Nova Southeastern University College of Optometry faculty and staff, in Fort Lauderdale, Florida. The commitment of Dr. Pizzimenti and the NOVA team enabled the successful implementation of the experimental plan through many hours dedicated to screening patient medical files and providing ophthalmic patient examinations.

I would also like to thank Kevin Moloney, Ji Soo Yi, and Young Sang Choi for donating a week out of their winter break to help me conduct the experiment at NOVA. Their commitment, positive attitudes, and eye for detail in the execution of this protocol made this experimental plan a reality and assured me of the validity in the data collected. In addition, Young Sang Choi is recognized for his outstanding work in the development of the software for the handheld PC.

I would also like to thank to Paula Edwards and Dr. Thitima Kongnakorn for their feedback and insight over the past year as I worked from my proposal to the completion of this thesis. Erin Kinzel is also acknowledged for her support with data entry and processing, which was invaluable to my timeline.

Family and friends have been an incredible support system during my academic endeavors. Their encouragement, love and support were instrumental in motivating me to see this work through to completion. I am especially thankful to my parents on so many levels. My Mom and Dad instilled in me the true value of education and hard work, which is demonstrated within this thesis. My brother and sister, Charles Emery and Jean Foster, have been indispensable sources of advice through the years and terrific role models for their younger sister.

Most importantly, I want to express my heartfelt appreciation for James, my husband and best friend. I am thankful for the active interest and excitement he has always shown for my work, including his development of the desktop software. His companionship, humor, patience and love have helped me to not lose sight of my goals along the way. I am excited for the road that lies ahead.

This dissertation research was made possible through funding awarded to Dr. Julie A. Jacko by the National Science Foundation (ITR-0106248;IIS-0308155 and BES-9896394) as well as an Intel Corporation Research Council Award (PI also J. A. Jacko). Additional funding was provided to V. Kathlene Leonard through a National Science Foundation IGERT Award granted to Dr. Marie Thursby.

TABLE OF CONTENTS

CHAPTER 5. RESULTS SUMMARY

CHAPTER 6. EXPLORING THE COMMERCIAL POTENTIAL OF ASSISTIVE TECHNOLOGY RESEARCH

LIST OF TABLES

LIST OF FIGURES

SUMMARY

This dissertation addresses the growing needs of a subset of computer users with visual impairments. The work considers the interactions of users who have been diagnosed with Age-related Macular Degeneration (AMD), the leading cause of blindness in adults 65 years and older. The investigation focused on the quantification of behaviors and strategies used when the visual sensory channel is compromised.

Participants diagnosed with AMD and age-matched controls without any ocular disease completed a series of visual search, icon selection and manipulation tasks on either a desktop or a handheld PC. Participants searched, selected and manipulated familiar playing card icons under varied icon set sizes, inter-icon spacing, icon sizes and auditory feedback. A comprehensive account of the interaction was made using a collection of efficiency, accuracy and information processing metrics. While all participants demonstrated a high rate for successful task completion, analyses revealed participants' overall task efficacy to be coupled with features of the interface and also strongly linked with measures of ocular health and personal factors.

The outcomes of this study contribute to a growing body of work which informs a framework of performance thresholds for critical graphical user interface interactions based on visual profile, interface features and supplemental non-visual cues. In addition, several notable results extend the existing

knowledge base of human computer interaction, aging and visual impairment including:

- The impact of auditory feedback on task interaction and information processing for visually impaired versus visually healthy older adults;
- The observed of use of the mouse pointer or stylus as means to direct attention during visual search and the implications of manual dexterity on visual search;
- The presence of speed accuracy trade-offs in handheld PC interaction performance for individuals based on their contrast sensitivity and near visual acuity;
- The shifting impact of increased icon spacing on visual search and movement times, versus its role in the accuracy of icon release;
- The utility for non-clinically acquired summaries of visual health to effectively predict performance decrements in handheld or desktop interaction;
- Emergent differences between handheld and desktop interaction and the most influential visual factors informing performance on each; and
- Empirical evidence that older adults, even with visual impairments can interact with small handheld displays, in spite of the size images.

An introduction to the problem and the detailed methodologies employed in are provided in Chapters 1 and 2. The results of the handheld and desktop PC experiments are separated into Chapter 3 and 4, respectively. Comparisons are drawn between the two platforms and the impact of the entire result set is explored in Chapter 5, discussing the implications for both research and design. This thesis concludes with a demonstration of the practical applicability of this type of research. This sample business plan provides substantiation for the potential impact for this research and illustrates the means through which universally accessible solutions can become part of a mainstream consumer technology marketplace.

CHAPTER 1

INTRODUCTION & BACKGROUND

1.1. Motivation

The ubiquity of information technologies in today's society generates a critical need for all citizens to be empowered to access and manipulate information electronically. Technical innovations, such as the graphical user interface (GUI), have made computing an integral part of life. The ability to interact with GUIs has emerged in recent years, as an essential component of work, family, recreation, and vital to accomplishing activities of daily living.

The advent of computers presented new opportunities for individuals with visual impairment to access digital information electronically, magnified, in Braille, or aurally. However, the introduction of GUIs, presenting information pictorially and symbolically, generated a digital divide for this population, and suddenly even the accessibility of electronic resources such as online catalogues proved complicated (Fortuin and Omata 2004). The success of GUIs is attributed to their exploitation of the visual sensory channel (Kline and Glinert 1995; Kline and Scialfa 1997; Jacko 1999; Jacko, Rosa, Scott, Pappas and Dixon 1999; Jacko 2000; Jacko, Barnard, Kongnakorn et al. 2004; Jacko, Moloney, Kongnakorn et al. 2005). GUIs facilitate the representation and manipulation of electronic information via visual metaphors including graphical icons with which the user interacts through a visual display and input device (e.g., a mouse or stylus).

The exclusive reliance of GUIs on the visual interaction paradigm can limit accessibility for anyone whose visual channel is compromised (Dix, Finlay, Abowd and Beale 1998). In the Using Statistics About Blindness and Low Vision Effectively (USABLE) Data Report #7 (Gerber and Kirchner 2001), data from the Census Bureau's 1999 Survey of Income and Program Participation (SIPP) was used to investigate Internet access and computer use by people in the US with visual impairments. Evidence of the digital divide was demonstrated in terms of access to technology and employment. In their publication titled, Foundation of a Conceptual Framework for Individuals with Disabilities, Jacko & Vitense note: "As the need for 'global information' grows, so does the variety of people requiring access to such information. As a result, a potentially large number of users may be disadvantaged with respect to accessing the diverse information available," (Jacko and Vitense 2001, p. 913).

It is estimated that nearly 20 million Americans have visual impairments resulting in low vision (Center on Aging Society 2002). This number is set to rise as aging baby boomers experience normal age-related changes to their functional vision (e.g., reduced visual acuity, presbyopia, contrast sensitivity, color sensitivity, depth perception, glare sensitivity) and ocular diseases associated with aging (e.g., Macular Degeneration, Diabetic Retinopathy, Glaucoma, and cataracts) (See Schieber 1994 for a review; Orr 1998). Some of these age-related ocular changes and diseases, such as acuity and cataracts, are correctable with lenses and/or surgery, while other conditions, such as Age-related Macular Degeneration (AMD), have no known remedy. Visual impairments encompass a range of functional limitations even in the presence of corrective lenses. Visual impairments can result in *severe limitations*, causing the individual to be *unable* to see words and letters in ordinary print, to *non-severe* visual impairments, causing the individual to have difficulty seeing print (Bailey and Hall 1989). This translates into difficulties performing other near vision tasks, such as using computers.

The interaction strategies and resulting interaction barriers for individuals with visual impairments in the past 15 years has received growing attention in an attempt to inform judicious, inclusive design for accessible information technology (e.g., Gaver 1989; Brewster, Wright and Edwards 1994; Jacko and Sears 1998; Jacko 1999; Jacko, Rosa et al. 1999; Fraser and Gutwin 2000; Jacko 2000; Arditi 2002; Jacko, Barreto, Scott et al. 2002; Craven 2003; e.g., American Foundation for the Blind 2004; Fortuin and Omata 2004; Jacko, Barnard et al. 2004; Jacko, Moloney et al. 2005). It is known, in terms of visual impairments, that user behavior is strongly influenced by the nature and amount of residual vision the user experiences in combination with computer interface characteristics. As an extreme example, a blind user without any functional vision will use fundamentally different coping skills to navigate an interface as compared to an individual with clouded vision due to cataracts (Jacko and Sears 1998). Despite this, while many assistive devices have come to market for individuals with visual impairments, they are typified by three underlying problems: 1) they present one size fits all solutions for a range of visual functionality; 2) they abandon the visual sense entirely, or only rely on the visual sensory channel; and 3) they do not accommodate changes in visual functionality over time.

The GUI interactions of older adult users who have visual impairments are of particular interest in the scope of this dissertation. In April, 2004, USA Today printed the headline: “Studies foresee increased vision loss among boomers,” (Kornblum 2004). The significance of research that examines this population is two-fold. First, the demographic make-up of the aging baby boomer population is a departure from the older adults of present day in terms of their familiarity with and perceived value of information technologies. By the year 2030, the population of Americans 65 and older will approach 70 million; currently 1 in 3 people experience vision-reducing eye disease by the age of 65 (Quillen 1999).

Computer experience is one trait that clearly sets the future older population apart from the current older adult population. According to the 2000 US Census, only 28% of adults 65 and older have home computer access compared to 51% for adults 55-64 and 65% for those 45-54 (Newburger 2001). Accordingly, as the baby boom population ages, it will be the first generation in which the majority of the members will already have significant computer experience when they reach the age of 65 (see Figure 1.1). The older generation is growing in terms of individuals who are comfortable and dependent on computers (Morrell 2002). This population has integrated information technology with GUIs into their daily lives; a visual sensory deficit could interfere with their ability to interact with the technologies, and compromise their independence.

Assistive technologies are being developed to extend the individual’s ability to live independently. For example, a critical component to obtaining important heath related information often involves using electronic health monitoring equipment, such as blood glucose monitors (American Foundation for the Blind 2004). If an individual can not fully receive the information presented by these devices, they may ultimately have to rely on others to do so in sacrifice of their independence. This population segment and their caregivers are therefore likely to proactively seek ways to extend their access to computers (in spite of visual or other age related impairments) and extend their independence through computers.

In 2003, the director of Microsoft's accessible technologies, Madelyn Bryant McIntire, in an Associated Press article discussing technology to improve quality of life for aging adults (Bergstein 2003), was quoted "If a boomer goes blind at 50, they're probably going to be far more motivated to have their PC remain a part of their life than the older person today. The voice of the boomers will come through loud and clear," (Bergstein 2003, p. 1).

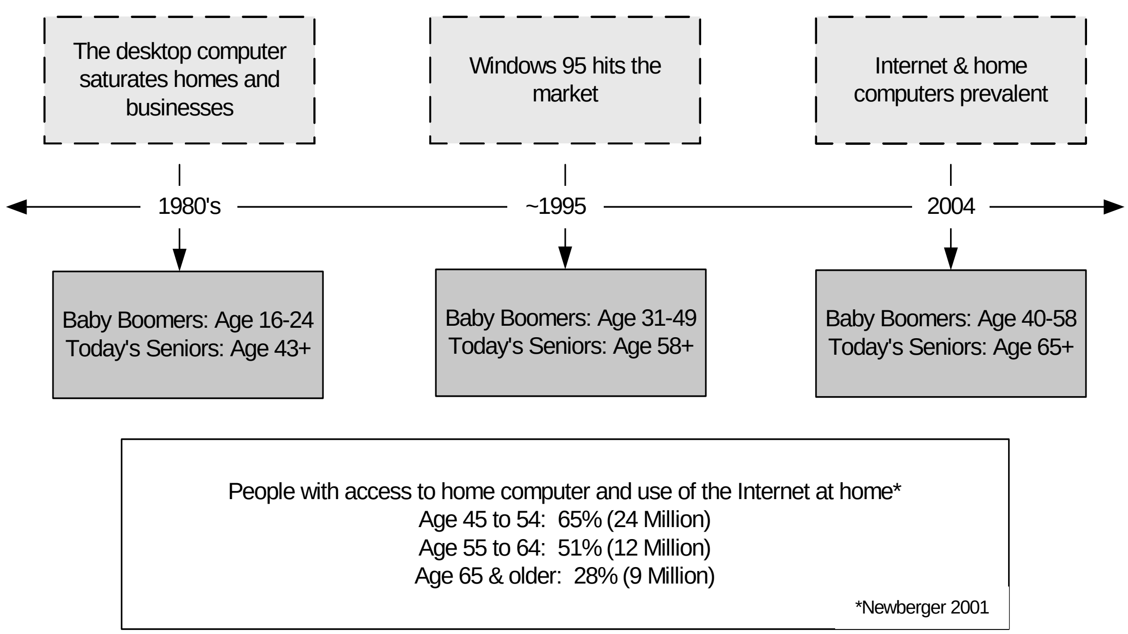


Figure 1.1. Timeline of aging Americans and developments in computing (Emery, Edwards, Jacko et al. 2003; Jacko, Emery, Edwards et al. 2004).

Table 1.1 provides the relative proportion of the US population who experience low vision disorders according to age bracket, and their projected growth through the year 2009. The projected growth of each segment is based on estimates taken from the 2000 US census. By 2009, the number of individuals with visual impairment is anticipated to be near to 25 million, with over half of those individuals being 45 and older. Without intervention, this rapidly expanding digital divide, emergent from the proliferation of technologies and rising number of aging adults, will have large scale societal impacts.

Table 1.1. Projected Number of Americans with low vision through 2009. Numbers based on the 2000 U.S. census and the percentage of individuals with low vision (Center on Aging Society 2002).

Population Age Segment	CAGR	2005	2006	2007	2008	2009
18-44 Year Olds	0.0%	6,046,000	6,070,789	6,095,679	6,120,671	6,145,766
45-64 Year Olds	8.1%	6,950,000	7,565,075	8,234,584	8,963,345	9,756,601
Age 65+	9.0%	6,311,000	6,937,051	7,625,206	8,381,626	9,213,083
Total	6.2%	19,307,000	20,572,915	21,955,469	23,465,642	25,115,450

*CAGR: Calculated Annual Growth Rate

Despite the apparent growing need for technology accessible to individuals with visual impairment, research has only started to address the interaction strategies and needs of this group. Literature that examines interactions beyond the traditional desktop PC is limited, yet the needs of users with visual impairments mandate investigation in a variety of contexts and platforms if this population is to keep up and profit from advances in ubiquitous technology. Hand held computers, mobile phones, kiosks, and personal computers need to be accessible to allow this population to sustain their independence. While some critical interactions for successful use are consistent between devices, many key characteristics of the interfaces themselves are inconsistent (e.g., screen size, graphic size, spacing, set size, input device).

This dissertation addressed the growing needs of a subset of computer users with visual impairments. The work considers the interactions of users who have been diagnosed with AMD, the leading cause of blindness in adults 65 years and older (Alberti, Richard and Sagerman 2000). The focus of the investigation is grounded in the considerations of behaviors and strategies associated with the performance of direct manipulation tasks that require visual search and iconic manipulation; interactions that have been recognized as critical to working in a GUI environment (Emery, Jacko, Kongnakorn et al. 2001). Assessments of the interactions are extracted as the users perform a series of drag and drops with icons. The effects of screen density (set size, icon size, icon spacing), and supplemental non-visual feedback are considered, in addition to the impact of the physical hardware with which the interaction occurs (desktop PC versus a handheld PC).

The exclusive focus of this dissertation on AMD is justified by the prevalence of this ocular disease in society and its clear impact on visual functioning that individuals rely on for normal daily activities (See the Literature Review in this document for details). Furthermore, the interaction needs of this user group are significant because those who acquire AMD are likely to experience the associated visual declines over time. There is no known cure and few surgical procedures for AMD (most only apply to individuals with a specific type of Macular Degeneration affecting 1 in 10 of the AMD population) (EyeMDLink.com 2002; VisionChannel.net no date). Additionally, these procedures are not corrective and cannot reverse the condition, per se, but rather were designed to slow and stabilize the condition (Alberti, Richard et al. 2000). For the time being, individuals with AMD manage the impact of this disease in activities of daily living with their own strategic coping skills, altering behaviors and making use of assistive devices to maintain their independence. Finally, unlike other ocular conditions, AMD almost never causes complete loss of vision (Quillen 1999). These individuals tend to rely on useful residual vision even as it changes over time (Owsley and Sloan 1990; Orr 1998; Jacko 1999).

This research builds upon seminal findings in the human-computer interaction (HCI) literature concerning individuals with visual impairments. It extends the research questions and results to an experiment with heightened ecological validity. Prior research has examined the impact of auditory feedback for this population in the context of a simple drag and drop task (single file to single folder) (Jacko, Moloney et al. 2005), while the present study took into account the interactions of drag and drop in the context of multiple icons and multiple folders, referred to herein as ‘complex.’ This research incorporates validated evaluation techniques and metrics attributed to HCI research within the mainstream user population, grounded on such established and recognized fundamentals.

Three components of this dissertation substantially advance the HCI and visual impairment knowledge base:

First, as stated, the drag and drop task employed in this study represents an increase in complexity from previous research on drag and drop with an AMD population. The interface used in the previous studies represented a simplified task environment (e.g., single file to single folder drag and drop) in order to isolate the effects of visual impairment to develop the fundamental, empirical knowledge of the interaction needs of users with visual impairments (Jacko, Moloney et al. 2005). This previous work guides the present research into user performance in more common and complex (multiple icons to multiple targets) task scenarios.

Secondly, the investigation of mobile computer use by individuals with AMD (or any other visual impairment) is a facet of HCI that remains largely uninvestigated, but is clearly warranted to alleviate the expanding digital divide.

Finally, while the impact of screen features (e.g., icon size, contrast, and colors) on the performance of users with visual impairments on a GUI-based visual search and targeting task have been examined (Jacko, Rosa et al. 1999; Jacko, Barreto, Chu et al. 2000; Jacko, Barreto, Marmet et al. 2000), the number of studies specifically addressing screen density (size, set size, and spacing) are limited. Furthermore, previous studies with this population focus primarily on screen density issues in terms of visual search and targeting only, and have yet to consider subsequent manipulation of the icons, such as dragging and dropping icons onto targets.

As will be demonstrated from the subsequent literature review, empirical work by Jacko and colleagues has established and contributed to a framework of interaction thresholds for individuals with visual impairments, as well as guidelines for accessible and universal design of GUIs. This framework is essentially the aggregation of results from these various studies, from which specific design recommendations can be derived, dependent on the task, technology, and user.

These studies considered visual screen parameters, multimodal feedback and the relevance of several measures to profile ocular functioning. As with the majority of innovative, exploratory research, a high level of experimental control was exercised in these studies in order to account for unexpected (and expected) confounds (Emery, Jacko, Sainfort and Yi in press). For example, in the study of multimodal feedback and the drag and drop task for users with AMD, the study focused on a simple task environment, that is a single file icon and single folder, to isolate the impact of the interaction apart from visual search and distracters (Jacko, Moloney et al. 2005). Notably, this investigation presents a substantial incremental increase in terms of task validity and context validity from the previous work in both the HCI and low vision research areas. That is, the drag and drop task performed is more representative of real-world task complexity (several icons and several targets), and the specific type of GUI (handheld PC versus desktop PCs) is to be examined. This work builds upon contributes to the growing framework established by Jacko et al., through a thorough examination of these interactions by individuals who have AMD.

1.2. Anticipated Outcomes

The fundamental goals of this study include the following activities that contribute to the body of knowledge considering HCI and visual impairment:

1. *Inform an increased level of judicious design of technologies so that they are more inclusive of the significantly growing population of older users with visual impairments.*

It has been asserted that bridging the digital divide for individuals with disabilities, such as visual impairments, requires the determination of 1) equipment needs; 2) Understanding of marketing and funding issues; and 3) addressing training needs. A critical step in achieving these goals is to ascertain the nature and extent of the problem (Gerber and Kirchner 2001).

2. *Investigate and compare interactions with handheld PCs and desktop PCs with an under-represented population of users.*

Very recent assertions claim that "despite the huge numbers, design-relevant data on visually impaired and elderly in general are rare…" (Fortuin and Omata 2004 p. 1). As stated previously, this research represents seminal work in the field of HCI and visual impairments in the context of mobile computing. Efforts to understand the impact of the physical hardware used in

HCI can help mitigate future barriers to ubiquitous computing for users who experience functional limitations due to visual impairments. The outcomes from this research yield design recommendations for more inclusive information technologies that can extend the independence of this population, a major concern with the anticipated growth surge of the older adult population segment. Within this dissertation, the commercialization of such research is discussed. A business plan detailing the commercialization, potential products and paths to market for this research are provided in Chapter 6.

3. *Further the knowledge base with respect to how visual function is linked to performance during use of desktop and handheld PCs in more complex and ecologically valid tasks.*

Scott, Feuer, and Jacko (2002) were the first to investigate the impact of visual acuity, contrast sensitivity, visual field, and color vision impairments in for a cohort of patients with AMD in the context of HCI. Performance of simple computer tasks involving mouse manipulations and icon selection in the environment of a desktop computer was investigated with measures of accuracy and reaction time. Results revealed that visual acuity and contrast sensitivity in the best eye, weighted average contrast sensitivity, and color vision defects were significantly associated with computer task accuracy. Visual acuity in the best eye, weighted average visual acuity, and color vision defects were significantly associated with performance speed. Visual function parameters associated with accuracy of computer task performance in a multiple regression model included weighted average contrast sensitivity (p=0.001), protan color vision defect (p=0.002), cataract grade in the better-seeing eye (p=0.036), and geographic atrophy outside the central macula (p=0.046). Visual function parameters and demographics associated with speed of computer task performance in a multiple regression model included color vision defects (deutan, p <0.001, and protan, p<0.001) and gender p=0.05) (Scott, Feuer, & Jacko, 2002). The present dissertation research extends the literature by further linking similar parameters of visual function with performance on more complex and ecologically representative tasks on both desktop and handheld PCs.

4*. Contribute methodologically to the field of HCI, particularly with respect to the involvement of subjects with visual impairments.*

This study extends significantly beyond previous efforts methodologically; considering more contextually representative GUI-tasks and extending the investigation beyond desktop computing to a handheld PC. Considering the limited amount of research in the area of visual impairment and HCI, the protocol, metrics, and analyses used in this research can serve as a model for future research involving individuals with visual impairments, and other user populations with diverse needs. This can inform further, meaningful exploration of related issues by other researchers.

5*. Address the needs of aging adults with computer experience, and their continued use of computers as they age.*

In the aging literature, the majority of existing research considers the training of older adults to provide them with computing skills. As previously mentioned, a distinguishing characteristic of the aging baby boomer population apart from the older adults of today is their computer experience, and the integral role of technology in their daily lives (Emery, Edwards et al. 2003; Jacko, Emery et al. 2004). To account for this, the current investigation incorporates computer experience in the inclusion criteria. This facilitates the application of these results to future, emergent GUIs and populations of older adults with computer experience.

The literature reviewed within this chapter reports how HCI researchers have traditionally examined iconic manipulation and drag and drop, as well as the current knowledge base in terms of non-visual, auditory feedback supplementing this type of GUI direct manipulation. The current state of the science in terms of visual impairments and HCI is also presented. Finally, the hypotheses that direct the dissertation are introduced.

1.3. Literature Review

This dissertation considers the unification of three established research domains: (1) Visual impairment and function (ophthalmic research), (2) Aging (gerontology), and (3) Human-Computer Interaction (HCI). While these three areas are well established in terms of significant theoretical and empirical work, the intersections of the three areas, illustrated by the Venn diagram in Figure 1.2, have not been as extensively considered. The * symbol in the center of the figure points to the domain addressed by this dissertation. The following background section reviews the three domains, drawing attention to those intersections that will be addressed by this dissertation.

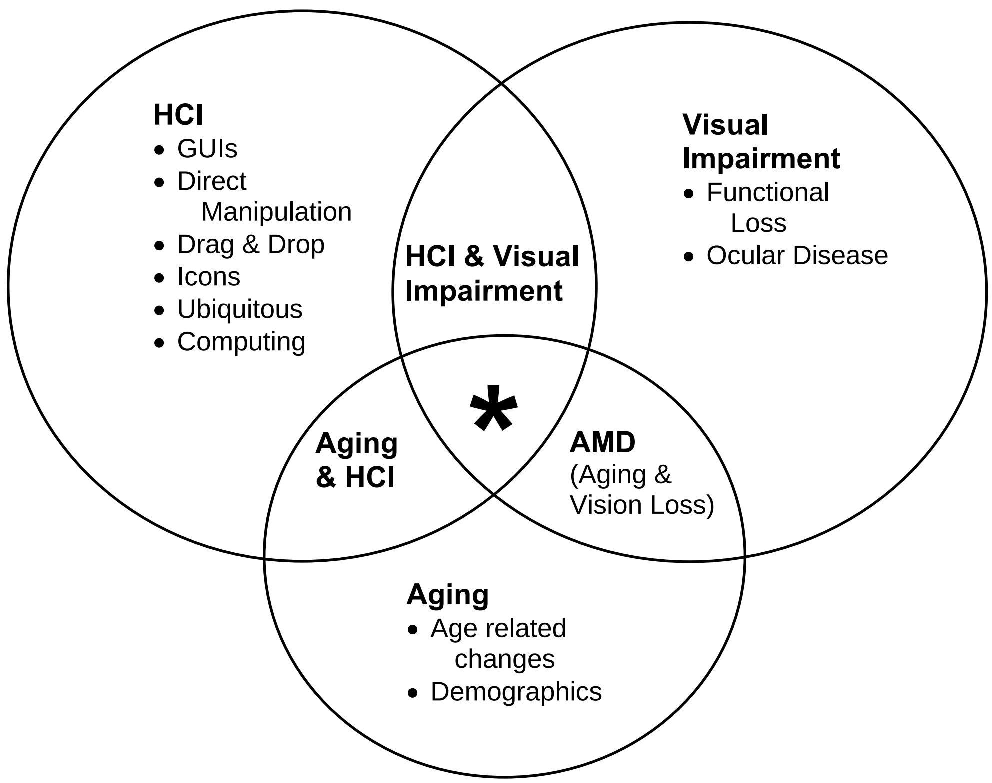


Figure 1.2. Venn diagram of research domains relevant to the dissertation. The various intersections represent the foci of this investigation, the * demarking the unique interdisciplinary union of topics which inform this work.

Visual Impairment

Visual impairment is defined by the American Foundation for the Blind (AFB) to encompass all degrees of vision loss, from slight visual field loss to total blindness

(Bailey and Hall 1989). In the scope of this dissertation, it is useful to differentiate amongst several terms that describe visual functioning beyond visual impairment, as it may be indicative of specific interaction strategies an individual is likely to employ to use GUIs. In addition to *visual impairment*, these terms include: *blindness* and *low vision*, illustrated in Figure 1.3.

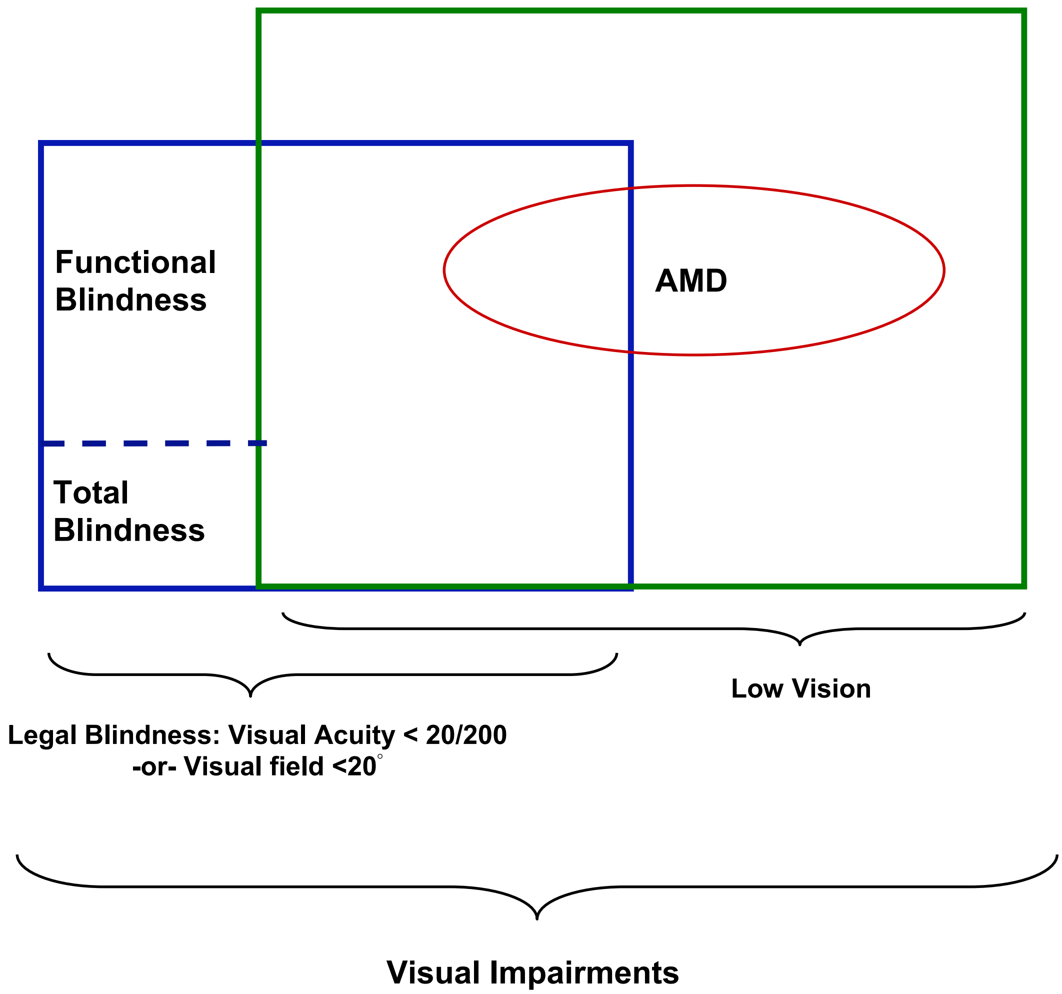


Figure 1.3. Schematic of relationships of types of visual dysfunction.

The term blindness has both legal and functional definitions. *Legal blindness* is a level of vision defined by government, and identifies individuals whose vision affords them certain benefits and services, as well as restrictions (such as driving). In the US, legal blindness is defined via visual acuity (the ability to resolve fine visual detail) and visual field (the physical space that is visible through the eyes). *Total blindness* refers to the complete loss of visual function, and *functional blindness* comprises individuals who have the ability to perceive light, but cannot resolve the shape or the source of the light. *Low vision* is visual loss that impedes tasks of daily living, while the ability to discriminate visual detail is still possible. Individuals classified as having low vision possess visual capabilities below what is considered normal (Biglan, Van Hasselt and Simon 1988; Jacko and Vitense 2001).

The AFB estimates that of the 1.3 million Americans who report legal blindness, 80% retain some residual vision (2004). It is also estimated that individuals with low vision presently outnumber completely blind individuals 3:1 (Newell & McGregor, 1997). Despite these statistics, research and development on assistive and universal technology solutions for individuals who are blind are more prevalent. Given the shear number of individuals with low vision, and projected growth rates, there is a clear need to investigate the nature of their interactions with information technologies. However, the disproportionate nature of solutions under development may be due in part due to an unanticipated paradox: *Providing access to users who have limited vision can be more challenging than providing access to users who do not see at all* (Arditi 2002).

Older adults, who experience vision loss, but retain some useful residual vision, are typically not willing to fully yield to these impairments. They instead develop coping strategies to use the visual capabilities they retain (Pelli, Robson and

Wilkins 1988; Ahn and Legge 1995; Brinker and Beek 1996; Jacko and Sears 1998). However, the level of visual impairment for this broad group of visually impaired individuals, and the relative effects that this impairment has on an individual's interaction with technologies, is quite diverse.

Visual function is most commonly assessed in terms of visual acuity (e.g., 20/20). In addition to visual acuity, other visual functions that may also be used to characterize low vision include: color perception, contrast sensitivity, eye movements, and visual fields (e.g., Jacko & Vitense, 2001; Jacko et al., 1999a, b) (For a complete reviews of types of vision loss and ocular abilities see Bailey and Hall 1989; Schieber 1994; Orr 1998). Ocular dysfunctions commonly correlated with aging are outlined in the following section.

Older Adults & Visual Impairment

Aging is synonymous with natural declines in a person's sensory abilities. As such, with age comes changes to the eye, including the retina and visual nervous system, that can impact functional vision (Schieber 1994). Additionally, older adults are more likely to acquire ocular conditions that can compromise visual functioning beyond normally anticipated changes, such as Macular Degeneration, Diabetic Retinopathy, and Cataracts. Age-related vision loss commonly impinges on the ability to complete near vision tasks, especially reading (Arditi 2004). An understanding of these functional declines provides direction for strategies aimed to mitigate the negative impact of these changes.

In her curriculum, *Issues in Aging and Vision,* Orr (1998) delineates the following 11 changes in ocular functioning that are viewed as typical of the aging process:

- Reduced visual acuity
- Reduced accommodation
- Reduction in visual field
- Decreased contrast sensitivity
- Color perception
- Floaters
- Dry eyes
- Increased need for light
- Difficulty with glare
- Dark/light adaptation
- Reduced depth perception

Table 1.2 provides a definition of each component of visual function, the impact it can have on computer use, and common clinically-based assessment techniques. (See Orr, 1998 and Schieber, 1994, for a complete review of the biological functioning and components of the ocular sensory system).

Table 1.2. Age-related visual functioning assessment method and impact on computer use.

Ocular Function	Definition	Age-related Change	Impact on Computer Use	Assessment Method
Visual Acuity	The ability to resolve fine detail in high contrast (Bailey and Hall 1989).	Reasonable declines in visual acuity typically occur beyond age 60, especially for conditions of low luminance and low contrast (Schieber 1994; Orr 1998).	The inability to resolve fine detail on a computer display has been shown to both increase the rate of errors, and time to complete direct manipulation tasks on GUIs (e.g., Jacko 1999, 2000).	Snellen Acuity is assessed via the ability to correctly identify visual characters at twenty feet compared to what a 'normal' individual can see at twenty feet (e.g. 20/20, 20/80, etc.) (Ferris, Kassoff and Bresnick 1982).
Accommodation	The ability of the lens of the eye to focus light rays onto the retina, or the focusing power as viewing distance changes (Sanders and McCormick 1993; Orr 1998).	Hardening of the lens and changes in controlling muscles leads to *presbyopia*; the decreased faculty to focus at close range. This is typically experienced after age 40 (Orr 1998).	If not corrected with lenses (bifocals, trifocals) eye fatigue or headaches may occur when reading or working off of computer display terminals (e.g., Orr 1998; Fraser and Gutwin 2000).	Measured with visual acuity scores (Sanders and McCormick 1993).
Contrast Sensitivity	The measure of how visible an object is before it is indistinguishable from the environment. It is a person's ability to detect small changes in brightness (Bailey and Hall 1989)	In the absence of ocular disease, neural differences in visual processing are attributed to much of the difference in ability between younger and older adults (See Schieber 1994 for a review).	To discern objects from their background, older adults benefit from higher contrast and sharper edges around objects and texts (e.g., Orr 1998; Jacko, Rosa et al. 1999).	Assessed using the Pelli-Robson chart (Pelli, Robson et al. 1988). This test assesses contrast sensitivity at different spatial frequencies.

Table 1.2. continued.

Ocular Function	Definition	Age Related Change	Impact on Computer Use	Assessment Method
Visual Field	The area of physical space that is visible through the eyes, the sensitivity of both the central and periphery of the retina. For normally sighted individuals, visual field is about 90 degrees on either side of the nose when the head is straight ahead (Bailey and Hall 1989).	Age-related decrements, have been observed to be, independent of optical factors, attributable to age-related neural loss (Owsley and Sloan 1990). Gradual loss is observed in the middle aged population, and accelerated visual field loss occurs after age 60 (Collin, Han and Korh 1988).	Blind spots in the visual field create barriers to a user's ability to systematically inspect a display (Bailey and Hall 1989). Visual field loss has been shown to increase the time needed for visual search and target selection in a drop down menu selection task. Restricted fields also make it impossible to survey the entire display at once (e.g., Fraser and Gutwin 2000; Jacko 2000).	Commonly measured with the Humphrey Visual Field Analyzer, which generates maps of a person's visual field, and measures an individual's ability to detect small spots of lights on a constantly illuminated background. Two key measures of the Humphrey Visual Field Analyzer are Mean Pattern Deviation (MD) and Pattern Standard Deviation (PSD) (Nelson-Quigg, Cello and Johnson 2000).
Dry Eyes	Eyes are susceptible to dryness, itching, burning and vision loss because a sufficient quantity of tears is not produced.	Older adults typically produce fewer tears, or tears of poor quality (Orr 1998).	Lengthy computer use has been associated with dry eyes (and subsequently blurred vision and eyestrain) (Lin 2004). Older adults are especially susceptible, and can interfere with the ability to work for a long period of time with a visual display.	There is no formal assessment of this condition. Artificial tear eye drops may mitigate the problem, and in the most severe cases, surgery (Orr 1998).

Table 1.2. continued.

Ocular Function	Definition	Age Related Change	Impact on Computer Use	Assessment Method
Color Perception	The ability of an individual to discriminate between colors.	While color blindness is often hereditary, a loss in ability to identify colors, especially those close in hue (blue-green), is common by the age of 60 (Schieber 1994; Orr 1998). Optical changes to the eye affect the illumination given to images on the retina, interfering with color perception (Owsley and Sloan 1990).	Color-coding of screen elements, especially blue-green coding, without effective illumination could lead to confusion for users with color deficits. Studies have shown that color perception is a predictor of performance with GUIs for individuals with visual impairments (e.g., Jacko, Rosa et al. 1999).	Assessed using the Farnsworth-Munsell 100 Hues Test. Participants arrange colored caps in an ordinal series extending between two anchored colored caps (Farnsworth 1947). The order of caps, as arranged by the individual is assessed to determine the presence and type of color confusion. This classification includes: protan, deutan, tritan, and non-congenital color deficiency.
Glare	Bright light reflecting from a surface that does not focus on the retina, but instead bounces around the eye. The stray light reduces the ability to resolve spatial detail and contrast (Schieber 1994; Orr 1998).	Changes to both the lens and retina sensitivity, inhibit older adults' ability to recover from, and tolerate glare (Schieber 1994; Orr 1998).	The dynamic nature of displays with respect to brightness and colors can impose performance decrements with diminished visibility attributable to glare.	While not typically part of the battery of clinical ocular tests administered, the Berkley Glare Tester (Bailey and Bullimore 1991) tests an individual's contrast sensitivity and visual acuity under conditions of high glare.

Table 1.2. continued.

Ocular Function	Definition	Age Related Change	Impact on Computer Use	Assessment Method
Dark/Light Adaptation	How well the eye adjusts itself between dark and light illumination conditions so the optical system can resolve small differences in luminance (Orr 1998).	The rate of dark adaptation significantly diminishes with age with respect to light sensitivity (See Schieber 1994 for further discussion).	The dynamic nature of displays with respect to brightness and colors can impose performance decrements with diminished visibility attributable to adaptation.	Measured by pupillary response, contrast sensitivity and visual acuity at different levels of luminance (See Schieber 1994 for further discussion).
Depth Perception	The ability to determine distance of objects in the environment.	Attributable to diminished contrast sensitivity, and accommodation, older adults may experience declines in depth perception (Schieber 1994; Orr 1998).	With respect to HCI, deficiencies in depth perception will have a significant impact in virtual reality applications.	Measured with the Worth-4 dot test. A patient wears glasses with one green lens and one red and looks at a target with 2 green dots, 1 red dot, and 1 white dot. Depending on the number and color of dots the patient sees, the examiner can determine whether if the vision of one eye is being suppressed.

Schieber (1994) reports a number of additional age-related functional visual characteristics that impact driving, which are also relevant to the use of GUIs and direct manipulation. A useful analogy is made between driving and using computers, as both require a significant amount of visual attention in order to orient and navigate to a final destination. Specifically, he describes the age-related changing attributes of eye movements, attention, and visual search as functions of the visual sensory system. The detection and orientation to events is a critical component to driving (Schieber 1994) and to other visually intensive tasks, such as interaction with GUIs. While driving and GUI use differ in terms of near and far vision requirements (driving requires both near and far vision, while GUI's typically require only near vision), they are analogous in terms of visual attention and motor skill requirements. In both tasks an individual must attend to task-relevant stimuli while rejecting the extraneous for efficient task performance.

Eye movements are a necessary component of processing information, especially in dynamically changing visual environments such as driving or GUIs. "Optimal spatial resolution depends on the ocular motor system's capacity to acquire, track, and image a visual target at or near the fovea," (Schieber 1994, p. 3). The fovea is the central portion of the retina with a high concentration of photoreceptors. Eye movements are therefore comprised of saccades and pursuit movements. Saccades are short, rapid movements of the eye with the purpose of centering visual information on the fovea. Pursuit eye movements are larger in nature and their purpose is to track moving targets (keeping the target in range of the fovea). An additional component of eye movements is the stationary periods, known as fixations that occur in between saccades, during which information may be processed (Sanders and McCormick 1993).

Schieber (1994) summarizes that age-related differences have been shown in saccadic eye movements, especially for the acquisition of targets in the peripheral vision. Findings point to a decreased accuracy of saccades for older adults, as more saccades were typically required for this population to fixate on a target in the peripheral vision, or in the context of searching complex visual scenes (see also Kline and Scialfa 1997; Lee, Legge and Oritz 2003). With respect to visual pursuit, age-related differences have been observed in the accuracy of tracking targets with higher velocities, further aggravated by the presence of distracting stimuli in the background or foreground. The perception of moving stimuli, for older adults, is both less effective and less efficient in tasks aimed at the detection of small target movement/change such as those found on dials and controls (Kline and Scialfa 1997). Furthermore, deficits in central, para-central, and peripheral visual field can pose different demands on vision, resulting in different search strategies and subsequent eye movements (Coeckelbergh, Cornelissen, Brouwer and Kooijam 2002).

Older adults, in research, consistently exhibit more difficulty with visual search tasks, especially when the number of items to be searched increases (Kline and Scialfa 1997). In studies of visual attention with older adults, tests generally report that the older population has longer reaction times due to the need to divide attention (Tun and Wingfield 1997; Ben-Shakhar 2001). Older adults are prone to difficulties ignoring “task irrelevant information,” (Schieber 1994 p. 17). Research also suggests that visual search is slower and less effective for older adults due to a shrinking of the “useful field of view” to which attention can be simultaneously allocated. The size of the useful field of view, for older adults, is especially susceptible to context related factors, such as complexity and cognitive task load (Schieber 1994).

Age-related Macular Degeneration (AMD)

In addition to the normally anticipated declines in vision, as highlighted in the previous section, the aging population is more susceptible to certain ocular diseases and conditions that can degrade visual functioning beyond what is normally anticipated. The four leading causes of visual impairment for aging adults include Macular Degeneration, Glaucoma, Diabetic Retinopathy, and cataracts (For a comprehensive review of each of these conditions, see Orr 1998; Quillen 1999).

A comprehensive review of the AMD was presented by Jacko and colleagues (2005), in their study of multimodal feedback as a solution to ocular disease-based performance decrements in the presence of functional visual loss. The leading cause of severe visual impairment in the aging population (individuals 65 years and older), AMD affects more than ten million Americans (Quillen 1999; National Eye Institute 2001; American Macular Degeneration Foundation 2002). This ocular condition is correlated with age; the majority of cases of Macular Degeneration observed in individuals 55 years of age and older (Quillen 1999; National Eye Institute 2001).

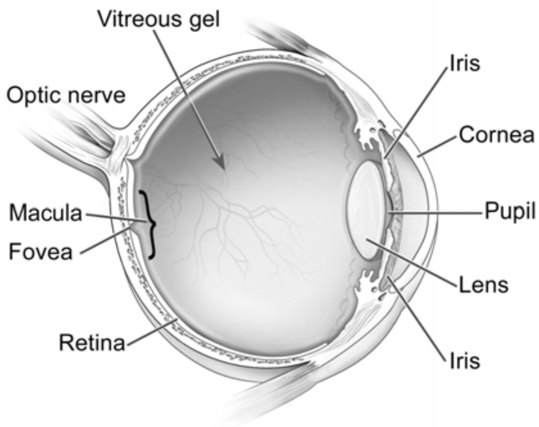


Figure 1.4. Anatomy of the eye, detailing the macula (Image Source: National Eye Institute and National Institute of Health, http://www.nih.gov).

AMD is a disease of the retina, affecting the macula or its central portion of the retina, illustrated in Figure 1.4. Roughly three millimeters in diameter, the macula is positioned near the optic nerve and is the densest locus of light sensitive cells (Macular Degeneration Partnership; Macular Degeneration Partnership).The macula is primarily responsible for central vision, fine detail vision, and color vision (Kaufman and Alm 2003). The progression of AMD entails deficits in central and high-resolution vision, which over time reduces sharp vision necessary to resolve objects and perform near vision tasks such as reading, driving, and using GUIs (Macular Degeneration Partnership; Orr 1998; American Macular Degeneration Foundation 2002). Table 1.3 presents an overview of the different types of AMD along with their associated ocular features and potential impact on visual functioning, and the images in Figure 1.5 illustrate this impact for varied levels of severity of AMD.

Figure 1.5. Progressive states of AMD (left to right); Blurred, distorted and occluded areas of the visual field are all typical impacts of AMD on visual functioning (photograph source: MD Foundation 2004).

Because computing technologies commonly employ GUIs that emphasize a visual feedback paradigm, users with visual impairments are at a distinct disadvantage (Farrell 1991; Jacko 2000). In addition, the direct manipulation paradigm employed in GUIs also requires the use of a pointing device, such as a mouse or stylus, which also generates visual processing demands such as visual attention and visual pursuit. The resources related to these visual functions are commonly lacking as a result of aging and/or age-related ocular conditions, such as AMD. The union of HCI (direct manipulation of GUIs) and visual impairments must be considered to generate potentially effective solutions.

Table 1.3. Understanding Age-related Macular Degeneration (Bailey and Hall 1989; Orr 1998).

AMD Classification	**Wet** *exudative/hemorrhagic*	**Dry** *nonexudative/atrophic*	
Statistics	• 10% of all AMD cases • 80% of Wet cases result in severe vision loss • Caused by the rapid growth of small blood vessels beneath the retina that leak blood and other fluid to form scar tissue	• 90% of AMD cases • 20% of nonexudative cases result in severe vision loss • Caused by thinning macula tissue triggered by drusen	
Visible Features	Macular Scar	Drusen	Geographic Atrophy
Description	Fibrosis developed from leaking of blood vessels creates a macular scar, caused by sub retinal fibro vascular proliferation, fluid, hemorrhage, and lipid exudate	Deposits of extra cellular material on the macula	Round, oval patches of atrophy on the retina that grow and coalesce
Impact on Visual Function	Severe vision loss to central vision central scotoma	Often no vision loss, or slightly blurred and distorted vision	Blurred, distorted vision, central scotoma, difficulty reading, driving, increased reliance on light and magnification

Visual Impairment & GUI Use

As previously stated, a disparity exists in the knowledge base of HCI for research aimed at individuals who are blind, versus individuals with low vision (Kline and Glinert 1995). The ranges in abilities of people who are visually impaired, but not blind, span a variety of functional abilities, attributable to residual vision, and personal attributes (such as gender, education and experience). Furthermore, it is estimated that individuals with low vision presently outnumber completely blind individuals 3:1 (Newell and McGregor 1997). Providing information technology access for users who have limited vision can be more challenging than providing access to users who do not see at all (Arditi 2002, 2003). To this end, concurrent developments have been made in terms of both the empirical research of interactions, and development of assistive and universal design strategies for individuals with visual impairments.

Strides have been made to determine the nature and extent of the problem of visual impairments (excluding blindness) and access to information technologies (Gerber and Kirchner 2001) A growing body of literature exists that exposes the detrimental effects of visual display elements on users with visual impairments, and mitigating design strategies, such as the inclusion of supplemental, multimodal feedback. This section provides a review of this literature, and concludes with observations on those needs yet unmet concerning the union of visual impairment and HCI.

At present, the majority of scientific literature concerning fundamental questions about computer use by people who are visually impaired and in some cases, elderly, is traced to work conducted by Dr. Julie Jacko and colleagues. Table 1.4 provides a summary of this literature, including the contextual elements and major findings of each study. The importance of understanding the specific details of a user’s impairment, in terms of their functional ability, as discussed by Jacko and Vitense (2001), are supported through the results of this body of work.

Table 1.4. Summary of research on GUI manipulations and visual impairment.

Citations	Cohort	Task	Independent Variables	Conclusions
(Jacko and Sears 1998; Jacko, Rosa et al. 1999; Jacko 2000; Jacko, Barreto, Marmet et al. 2000)	Individuals with ocular diseases that result in severe visual impairment	Visual search and selection	• Icon size • Set size • Background color • Visual profiles	• Visual acuity, contrast sensitivity, color perception and visual field are significant predictors of performance. • Icon size, set size, and background color significantly influence performance. • The increased time required by individuals with visual impairments to search is not due to delayed engagement, but to time spent in active search.
(Jacko, Scott, Barreto et al. 2001; Jacko, Barreto et al. 2002; Scott, Feuer and Jacko 2002, 2002)	Individuals with AMD	Visual search and selection	• Icon size • Set size • Background color • Visual profiles	• Visual acuity, contrast sensitivity, weighted average visual acuity, contrast sensitivity and color perception deficits were found to be significantly associated with the performance of this search tasks for users with AMD. • Weighted average contrast sensitivity was the most sensitive indicator of performance. • Analyzing eye movements confirmed there were differences due to AMD in the visual search and the interface features of icon size, background color, and set size. • Changing key screen features such as icon size, background color, and set size can improve performance for individuals with AMD.
(Jacko, Barreto, Marmet et al. 2000)	Individuals with AMD	Cursor Movement	• Icon size • Set size • Background color • Visual profiles	• Cursor movement time and velocity were significantly worse for individuals with AMD, and worsened in conjunction with visual acuity.

Table 1.4. continued.

Citations	Cohort	Task	Independent Variables	Conclusions
(Jacko, Scott, Sainfort et al. 2002; Jacko, Scott, Sainfort et al. 2003; Jacko, Barnard et al. 2004; Jacko, Moloney et al. 2005)	Individuals with AMD	Drag and Drop	• Supplemental multimodal feedback (haptic, auditory, and visual) • Visual profiles	• Performance improvements were observed for both visually healthy and AMD users due to the implementation of non-visual/multimodal feedback. • Significant differences between groups of different visual acuity were observed, including task time, feedback exposure times, and frequency of errors. • The performance gains for the utilization of non-visual, multimodal feedback were greater in magnitude for users with AMD. • The presence of AMD significantly inhibited user performance, independent of other ocular functions (e.g., acuity, contrast sensitivity, color perception). • Individuals with AMD demonstrate gross head and fine eye movements during task performance compensating for central vision loss. • When administered VFQ 22, a self assessment of the impact of visual impairment on daily living, participants with AMD rated their general vision, role difficulties and mental health significantly lower than the cohort of individuals without AMD, even though functional visual characteristics were not different between the two.
(Edwards, Barnard, Emery et al. 2004; Edwards, Barnard, Leonard et al. 2005; Jacko, Barnard, Yi et al. in press)	Individuals with Diabetic Retinopathy	Drop Down Menu Selection	• Windows Accessibility Setting features (e.g., size & contrast) & Multimodal feedback • Visual Profiles	• Multimodal feedback was found less effective than visual enhancements to the selection of items from drop down menus. Ocular factors such as acuity, contrast sensitivity, and visual field were found to impact performance.

Jacko and colleagues have illustrated the detrimental effects of the visual processing demands imposed on users with visual impairments by technologies employing GUIs (e.g., Jacko and Sears 1998; e.g., Jacko, Rosa et al. 1999; Jacko 2000; Jacko, Barreto, Marmet et al. 2000), with significant attention to users with AMD (Jacko, Scott et al. 2001; Jacko, Barreto et al. 2002; Jacko, Scott et al. 2002; Scott, Feuer et al. 2002, 2002; Jacko, Scott et al. 2003; Jacko, Barnard et al. 2004; Jacko, Moloney et al. 2005). As summarized by Table 1.4, these studies have addressed the relative performance of a cohort of users with visual impairments and a cohort of age-matched controls without ocular dysfunction in the context of:

1) The visually rigorous task of icon search and selection in the presence of distracters (Jacko 1999; Jacko, Rosa et al. 1999; Jacko, Barreto, Marmet et al. 2000; Jacko, Scott et al. 2001; Jacko, Barreto et al. 2002; Scott, Feuer et al. 2002);

2) Cursor movement (Jacko, Barreto, Marmet et al. 2000); and

3) The direct manipulation of drag and drop in the absence of distracters (Jacko, Scott et al. 2002; Jacko, Scott et al. 2003; Jacko, Barnard et al. 2004; Jacko, Moloney et al. 2005) and the identification and selection of targets in a drop down menu with distracters, (Edwards, Barnard, Emery et al. 2004; Edwards, Barnard et al. 2005; Jacko, Barnard et al. in press).

In those studies concerning target selection tasks in the presence of distracters, participants were instructed to select an identified target icon from a set of two or more icons. The icons employed in these experiments were those analogous to Microsoft® Word: Print, Paste, Save, Copy, New, and Open, due to their identifiable nature of the icon by users of Microsoft® Word® as recommended by Sears and colleagues (Sears, Jacko, Brewer and Robelo 1998). Five different icon sizes were employed, 9.2 mm, 14.6 mm, 23.2 mm, 36.8 mm and 58.3 mm, which corresponded to the size of letters on the Bailey-Lovie acuity chart. Five different background colors were manipulated, using black, white, blue, red, and green. Performance was assessed in terms of target identification reaction time, accuracy of selection, and eye movement analysis. Results revealed that several aspects of visual functioning have a significant impact on task performance, including visual acuity, contrast sensitivity, color perception, and visual field.

Finally, it was concluded that there is a point of diminishing return associated with increasing icon size. As icon size increases, it can actually become less visible for certain visual dysfunctions particularly for losses in visual field. Effectively, portions of an enlarged icon may be occluded by blind spots in the visual field, calling for studies with more subjects and specific diagnoses to develop more universally effective guidelines.

Contrast sensitivity and color perception were found to be significant when predicting performance time, while visual acuity, contrast sensitivity, visual field, and color perception were all found to have a significant effect on icon selection accuracy. These studies led to the conclusion that the functional vision profiles should be accounted for in understanding the performance thresholds for GUI use by individuals with visual impairments.

For a similar task, the source of time delays in task performance for individuals with visual impairments were considered in another study by Jacko and colleagues (Jacko, Barreto, Marmet et al. 2000). Using electroencephalogram signals generated during the execution of the selection task, the source of additional time required was determined to be a result of time spent in active search once the visual cortex has already been engaged. The differences in focal attention suggest that individuals with visual impairments may have more difficulty making the comparative judgments of discrimination between stimuli that occur during prolonged focal attention. This suggests that they have relatively more trouble with the extended processing of visual stimuli. Neither the nature of their eye movements in the directed search, nor their processing of visual information is optimized. As a result, the nature of information processing of the individuals with visual impairment likely contributes to the performance differential between this group and a visually healthy, age-matched control group.

In terms of those studies on drag and drop manipulations and the inclusion of multimodal, non-visual feedback performance was compared between users with the ocular condition of AMD, and within those who had AMD, different levels of visual acuity. For this simple drag and drop computer task (single file to single folder) both users who are visually healthy and users with visual impairments performed worst when presented with only visual feedback. Performance was assessed in terms of total task time, the amount of time the user was exposed to the feedback before successfully dropping the file into the folder, and the number of missed opportunities (times they could have successfully released the file into the folder, but continued mouse movement to move out of position). The inclusion of auditory and haptic feedback resulted in some performance improvements (although inconsistent between the visually healthy group and those with AMD). Non-visual feedback, presented by itself and in addition to visual feedback, resulted in strong performance improvements compared to performance when visual feedback was presented alone.

One study of drag and drop performance and multimodal feedback highlighted that the exclusive reliance on visual interaction paradigms clearly neglects the potential sensory deficits of users with visual impairments and does not truly support interaction via other sensory channels (Jacko, Moloney et al. 2005). Sensory channels go largely untapped, because most efforts of computer-based information presentation have focused on the visual display of information (e.g. text, graphics, animation), potentially overwhelming those with limited visio-spatial capabilities (Stanney et al., 2003). Interfaces which employ multimodal feedback have been shown to improve user interaction with computers by utilizing multiple perceptual processes, allowing for enhanced information processing through parallel sensory channels (Brewster, Wright et al. 1994; Vitense, Jacko and Emery 2002, 2003). For the simple drag and drop task, the effectiveness of non-visual feedback increased relative to severe visual acuity loss (e.g., Jacko, Scott et al. 2002).

However, while multimodal feedback has been shown to augment the perceptual processes and induce performance gains for individuals with visual impairments performing computer-based tasks, this is not universally true. This was the case for a comparative study of the Windows® accessibility settings and multimodal feedback applied to a menu selection task (Edwards, Barnard et al. 2004; Edwards, Barnard et al. 2005). The impact of multimodal feedback on performance in this task was insignificant. That said, the effectiveness of strategies for improving access to GUI technologies for users with visual impairments is also dependent on task characteristics. Table 1.5 aggregates these studies by Jacko and colleagues to effectively illustrate where the deficiencies are in the visual function and HCI research, and where this dissertation adds incremental value to the knowledge base.

Table 1.5. Illustration of topics covered in previous visual impairment and HCI by Jacko and colleagues; a shaded cell in the table indicates that the topic had been addressed by those studies listed in the left-hand column.

	Visual Impairment			Age		Task Phases					Com-plexity		Independent Variables						
Citation	Visually Healthy	AMD	DR	General Population	Older Adults	Visual Search	Object Selection	Cursor Movement	Drag & Drop	Menu Selection	Simple	Complex	Icon Size	Set Size	Background Color	Clinicall Visual Profiles	Non-ClinicalVisual Profiles	Multimodal Feedback	Text Size
Jacko 2000; Jacko Barretto, Scott, Rosa & Pappas, 2000; Jacko et al., 1999; Jacko & Sears 1998		■	■	■	■	■	■				■	■	■	■	■				
Jacko, Barreto et al., 2002; Jacko et al., 2001; Scott, Feuer & Jacko 2002a Scott Feuer & Jacko 2002b		■			■	■	■				■	■	■	■	■	■			
Jacko et. al, 2000		■			■			■				■	■	■	■	■		■	
Vitense, Jacko and Emery, 2002; Vitense, Jacko and Emery, 2003	■			■					■			■						■	
Jacko et al., 2002c; Jacko et al., 2003; Jacko et al., 2004a Jacko et al., 2004b; Jacko et al, 2005 Leonard et. al, 2005		■			■				■		■					■	■	■	
Emery et al, 2004; Jacko et al. 2004c	■				■				■		■							■	
Edwards et al., 2004; Edwards et al., 2005 Jacko et al, in press			■	■	■	■				■		■	■		■	■	■	■	■

Additional research and development has contributed to this subject area in the past 10 years. A review of the research and common solutions provided for the access of digital information for individuals with visual impairments (who maintain residual vision) yields that the majority of knowledge (with rare exception) resides in (1) the magnification of screen elements (Fraser and Gutwin 2000); and (2) the accessibility of text (Craven 2003). Comparatively, less has been accomplished in terms of critical aspects of the graphical user interface, either empirical or development based.

In their seminal work, Kline and Glinert (1995) presented UnWindows V1, a set of interface tools to support selective magnification of window area, and tracking the location of the mouse pointer on the display screen. The authors note that "Magnification is one method commonly employed to help low vision users deal with the small type fonts, illustrations, and icons present in much of today's printed media and computer displays," (Kline and Glinert 1995 p. 2). Key components of the UnWindows system included: 1) a dynamic magnifier to compensate for the loss of global context imposed by static magnification and changing display content; and 2) Visual and aural feedback to aid the users in locating the mouse pointer. Kline and Glinert placed emphasis on the problematic nature of visual tracking in the presence of a screen densely populated with icons and windows. Interestingly, they received mixed reaction to their interface by users with and without visual impairment, especially in terms of the auditory feedback provided whenever the mouse pointer entered a new window (users found this annoying). And while no formal empirical testing was performed in relation to UnWindows, questions surface as to the effectiveness of non-visual, multimodal feedback in a complex display.

Fraser and Gutwin (Fraser and Gutwin 2000) discuss the impasses imposed by the mouse pointer to direct manipulation for individuals with low vision. Having low vision creates barriers to distinguishing fine details of iconic screen targets, as well as tracking the highly dynamic nature of the pointer used to manipulate these icons. The authors attribute the difficulty in manipulating objects with the pointer mainly to reduced visual acuity, and constrained visual field on the basis of four dimensions:

- *Mode:* The sensory channel through which assistance is provided to the user;
- *Stage*: The phases of targeting supported by the pointing solution, including a) locating the pointer, b) moving the pointer towards the target and c) acquisition of the target;
- *Dependence*: How the pointing solution, interface, and the onscreen pointer are interconnected; interface dependent or independent; and
- *Pervasiveness*: The balance of availability of the assistance and intrusiveness on the goals of the task; fixed, selective, consistent, and requested assistance.

While these four dimensions are intended to evaluate the effectiveness of assistive mouse pointers, they also have a bearing on the effectiveness most direct manipulations with GUIs employing the Windows-Icon-Menu-Pointer (i.e. *WIMP*) interaction paradigm. In their review of assistive technologies for the visually impaired the authors conclude that magnification is not appropriate for users with severely limited visual fields (Fraser and Gutwin 2000).

A usability review of currently available technologies for the conversion of GUI technology for use by individuals who were blind or possessed low vision, the ability of magnification, synthetic speech, and Braille were reviewed for their ability provide the respective users 100% access to GUI's on nine test areas (Becker and Lundman 1998). These tests included:

- Installing and configuring the device/software;
- Uninstalling the device/software;
- Performance reliability/stability;
- Program manager to read and manipulate windows, menus and icons;
- Word processing based tasks such as opening and saving files, reading and editing text, text attributes, and toolbars;
- Spreadsheet tasks such as reading cells, tables figures, and editing data and formulas;
- Internet use, including dialing up, accessing World Wide Web pages, navigating with link buttons, sending e-mail and reading graphics;
- Screen searching, e.g., searching for characters, strings, formats and icons; and
- Operating start menu, exploring and controlling settings.

For the assessment of seven magnification programs (synthetic speech and Braille displays are not relevant to individuals with visual impairments who have residual vision), the evaluators comprised of a system engineer, ergonomic engineer, computer science expert, and three individuals with visual impairments.

The use of magnification, as a strategy to afford access to GUIs proved somewhat successful, providing 89% or higher access to GUIs, except for Internet use (84% access) and screen searching (0%) (see Fortuin and Omata 2004 for a review of this study.) It was concluded that the essential problem for the design of interactive systems for users with visual impairments is to 1) to determine what the users need and 2) how to represent the requested information based on key psychological and physical attributes of the user. The result of ineffective assistive technologies is a lack of usable contextual cues for the users to provide feedback in the case of errors; and this translates to large amounts of imposed workload on the user and frustration (Fortuin and Omata 2004).

In a case study on an English teacher who was having difficulty reading student papers, typing and proofreading, Whittaker and Young (Whittaker 1998) discovered that magnification was not affording optimal performance. Typically, the authors found that users with visual acuity of 20/40 or better would respond well to simple optical magnification. The authors investigated other visual functioning to find that the individuals had severely diminished contrast sensitivity (13% contrast threshold, with 2% representing normal sensitivity). Furthermore, this individual's visual field was 20 degrees horizontally (180 degrees is normal). Magnification was likely reducing the number of letters viewable simultaneously in the presence of scotoma within the visual field. The author warns that magnifiers and large monitors are not always the most effective solution for users with impaired vision.

Arditi addressed the reading difficulties of individuals with low vision (Arditi 2004). According to the author, successfully overcoming this difficulty is accomplished through the exploitation of remaining vision. The easiest way to do this is through magnification, but as shown in this study, it is not a one size fits all solution. Several parameters of the font, including height, stroke, spacing, and serif size, must be selected in a combination that best suits a given user. Arditi presents the prototype and initial user testing of computer-based software that lets a user customize fonts for maximized legibility. Those users studied were able to adjust font to a usable, legible level, to positively impact reading times and the reading acuity.

Ludi considered the animation of icons as a means to reconstruct visual cues for computer users with visual impairments (Ludi 2000). Her research questions considered: 1) the optimal size for animated icons for partially sighted users; and 2) perceived differences in size between static and animated icons. However, the results of this study (as presented in a poster session) have gone largely unpublished.

An in-depth review of accessibility tools aimed at improving interactions of computer users with low vision informed the design of *MouseLupe* (Silva, Regina and Bellon 2002). *MouseLupe* simulates a magnifying glass, enabling users to magnify select portions of text or display graphics, inspired by the problematic nature of screen magnification software. The authors suggest that magnification improves the readability of smaller text, but occludes the visible area of the document.

Furthermore, graphics that contain text (like most icons), a critical element of the graphical user interface, when enlarged, are difficult to read (see Silva, Regina et al. 2002 for a comprehensive review of magnification tools).

Several scholars have considered the effect of low vision on web browsing (Harper, Goble and Stevens 2001; Silva, Regina et al. 2002; Arditi 2003; Craven 2003). Harper, Gobel and Stevens address this problem in terms of “Web Mobility,” (Harper, Goble et al. 2001). These authors provide guidelines for movement through and around complex hypermedia environments, such as the web, for users with visual impairments. The problem, according to these authors, is that low vision inhibits the individuals’ ability to efficiently assimilate page structure and visual cues that lead to the following problems:

- Failure to get a feel for the content on the website
- Failure to have a sense for the magnitude of the display or where in a website the interaction takes place
- Disorientation
- Obstacles and distracters such as spacer images, tables, and large images
- Too much complex detail that cannot be resolved
- Frustration

Harper and Gobel emphasize that the differences in orientation, navigation, travel, and mobility of visually impaired versus sighted individuals should be considered in the design of technology because there are differences in the mental map and cognitive processes that occur across the spectrum of visual ability (Harper, Goble et al. 2001).

Arditi (2003) has observed the problems of Web browsing in terms of the allocation of screen space resources. According to the author conflicts arise in the implementation of Web browsing solutions for individuals with low vision including: 1) high magnification requirements; 2) variable typography color, size and contrasts of the content presented; 3) embedded text messages to augment Web images; and 4) accessible Web browsing controls (icons, buttons, menus). The author presents a novel approach for effectively using screen resources, providing evidence that the strategic layout of a display is a critical factor to successful interaction. The layout of screen elements was interpreted as more critical than magnification of the screen elements.

Craven (2003) questioned the accessibility of electronic library resources on the World Wide Web for individuals with visual impairments. The results of her study with 20 sighted and 20 visually impaired users revealed the browsing times of those individuals with visual impairments were significantly greater, depending on the design of the Web site (layout complexity and distracters). Navigation time for the group of users with visual impairments was significantly longer due to visual functioning, but also due to artifacts of assistive technology use in navigation (magnification and screen readers).

HCI research and development activities targeted at this population with visual impairments have yet to fully explore contexts beyond desktop computing. This includes key features of ubiquitous computing, such as mobile computing (including wireless technology) (Baker 2004), and kiosks (Vanderheiden 2004). Yet, if these areas of GUI interaction are not attended to, the digital divide imposed on users with visual impairments will only grow.

Researchers have only recently begun to ask questions concerning the use of mobile, wireless technologies by users with limited abilities such as visual impairment. Mobile computing introduces challenges to HCI, providing access to “powerful computing services and resources through small interfaces, which have tiny visual displays, poor audio facilities, limited input techniques,” (Dunlop and Brewster 2002). Interactions with mobile computers are also susceptible to the affects of context; varying tasks, environments, and even users (Barnard, Yi, Jacko and Sears 2004, 2004) Situationally-induced impairments (SII), a term coined by Sears, Lin, Jacko and Xiao (2003), are the extraneous demands imposed on the user by the context of use that interfere with optimal task performance. SIIs introduce barriers to the completion of a task.

Users with visual impairments who wish to use mobile computing technologies, such as cell phones and handheld PCs, are likely to encounter these SII in addition to barriers to interaction imposed by their functional vision, also known as disability induced impairments (DII) (Sears, Lin et al. 2003). Intuitively, some interaction effects are anticipated between SIIs and DIIs, although the precise effects and magnitudes have not yet received attention in the research community. In fact the consideration of mobile computing for individuals with any limitation in physical, sensory, or mental ability has only recently received attention on the forefront of the research community’s agenda. Barriers to the use of wireless technology by individuals with disabilities have been categorized as 1) economic in nature; 2) awareness of and proficiency with wireless technologies; and 3) compatibility of the device with other assistive technologies (McNeil and Griffin 2002). In spite of these barriers, it is asserted that individuals with disabilities can find great uses for these devices, such as way-finding, cognitive reminders, and communication devices (McNeil and Griffin 2002).

Fruchterman (2003) anticipates that through research and development, the cell phone will actually become the “Swiss army knife” for individuals with disabilities, including visual impairments. The author agrees with McNeil & Griffin (2002), that the major barriers to the use of mobile technologies by individuals with visual impairments include price and complexity, and this effect has been observed already in the cell phone market. For example, Fruchterman envisions digital camera cell phones, in the relatively near future, will be able to orally describe to the user key features of the images captured by the camera.

In an article by the *Associated Press* titled, Technology for Better Living (Bergstein 2003), a fifty-year-old woman with a severe visual impairment noted that technology solutions that go beyond text enlargement are growing, but in terms of mobile technology, the small size of cell phones and associate controls can be “a nightmare.” Smith-Jackson and colleagues (2003) used semi-structured interviews and focus groups to derive accessible design requirements for cell phones, which matched six of the seven principles of universal design (Connell, Jones, Mace et al. 1997).

There is a negligible amount of previous work which considers the use of mobile technology by individuals with visual impairments. It tends to be highly subjective, anecdotal and speculative. Researchers are at present engaged in the formulation of the problem, and research extending beyond subjective usability analyses is uncommon. Limited research is published reporting objective studies concerning the interactions of users who have visual impairments with mobile technologies. One exception that surfaced in a review of literature objectively investigated a prototype of a multimodal handheld PC that integrated tactile and auditory feedback (Amar, Dow, Gordan, Hamid and Sellers 2003). However, this was an extended abstract for a poster presented at the CHI 2003 conference, and hence limited in scope, and its mix of subjective and objective usability testing metrics (e.g., time, user interviews). The authors did conclude that "the means for providing necessary enabler for the visually impaired is not simple" (Amar, Dow et al. 2003, p. 919).

Tables 1.6 and 1.7 summarize the outcomes of these additional HCI and visual impairment research. From this review of HCI research and development that targets individuals with visual impairments; an agenda of unmet needs can be constructed. At the forefront of this is a growing lack of correspondence between empirical work, assistive applications development and mainstream HCI research and development. For example, the majority of solutions for improving user performance and providing access for individuals with visual impairment with GUIs focus on 1) text; 2) augmentations to the visual component (usually through magnification); and 3) desktop computers.

Table 1.6. Summary of additional research on HCI and visual function: a shaded cell indicates that the topic has been investigated by the citation listed on the left.

	Magnification	Direct Manipulation	Multi-Modality	Icon Quality	Layout	Web Navigation	Screen Readers
Mouse Lupe (Silva, Regina and Bellon 2002)	■						
Framework for assistive pointers for low vision (Fraser & Gutwin 2000)	■	■					
Unwindows VI (R.L. Kline & Glinert 2000)	■	■	■				
Animated icons for low vision computer users (Ludi 2000)			■	■			
Access to electronic library resources (Craven 2003)	■				■		■
Web mobility - movement through complex hypermedia (Harper, Gobel & Stevens 2001)						■	
Allocation of screen space for web browsing (Arditi 2003)	■				■	■	

Table 1.7. Major conclusions emergent from additional HCI research on visual impairment.

Investigation	Conclusions
Mouse Lupe (Silva, Regina and Bellon 2002)	Graphics that contain text are difficult to read when enlarged (e.g., icons).
Framework for assistive pointers for low vision (Fraser & Gutwin 2000)	Magnification is not appropriate for users with severely limited visual fields.
Unwindows VI (R.L. Kline & Glinert 2000)	Intelligent magnification is useful, but auditory feedback in complex task scenarios is reported to be annoying.
Animated icons for low vision computer users (Ludi 2000)	Differences were observed between perceived size of static versus animated icons for partially sighted users.
Access to electronic library resources (Craven 2003)	Web sites with higher layout complexity and distracters imposed longer browsing times, attributed to the use of magnification and screen readers.
Web mobility - movement through complex hypermedia (Harper, Gobel & Stevens 2001)	Low vision inhibits an individuals' ability to assimilate page structure and visual cues; Orientation, navigation, travel and mobility are required actions for all users and differ as a result of mental maps and cognitive processes that are closely linked to visual ability.
Allocation of screen space for web browsing (Arditi 2003)	Conflicts exist between the solutions for web browsing for individuals with low vision; Simplistic layout is a more critical feature than magnification for web browsing by users with visual impairments.

The bulk of published empirical work, explores the impact of both visual and non-visual augmentations to the interface, but focuses exclusively on interactions in the context of desktop personal computers. As previously mentioned, this foundational work has exercised significant control in laboratory-based experiments. For example, work with drag and drop has been accomplished with a single file and single folder (Jacko, Scott et al. 2003), while work with a visually healthy population has examined a higher level of task complexity (multiple files and folders) (Brewster

1998; Vitense, Jacko et al. 2002, 2003). Furthermore, as will be discussed in the following section, several of the advancements in the evaluation of HCI for mainstream GUI technologies have yet to be applied to divergent user populations, such as users with visual impairments. This includes the evaluation of fundamental interaction paradigms and interface components, particularly icons.

Direct Manipulation Tasks & GUIs

GUI-based computers and technologies are widespread in the consumer market. As stated, these interfaces rely a great deal on the graphical symbols and icons as both visual elements of the screen and tools for interaction. These screen elements serve as the representation of low-level computer functionality that users can manipulate through higher-level actions without having to use a complex syntax or direct programming language (Shneiderman 1998). The graphics facilitate control of the complex computing functionality through the common actions of touching screens, pressing buttons, manipulating icons and objects, and moving the cursor on the screen with an input device, like a mouse or a stylus (Jacko, Rosa et al. 1999). This execution of complex computational functionality via the simple manipulation of screen element is called direct manipulation (Shneiderman 1998), a “visual interface which emphasizes eye-hand coordination skills as a prime requisite for successful and efficient interaction,” (Eason, Johnson and Fairclough 1991, p. 116). While direct manipulation affords individuals access to computing without knowledge of syntax or coding, it is especially prone to design flaws; poor design, slow implementation, or inadequate functionality can negate any of the advantages created by the paradigm (Shneiderman 1998).

Emery et al. (2001) introduced a taxonomy of critical interactions for GUIs. The intent was to guide evaluators in choosing key interaction scenarios relevant across different GUI-based technologies and users. Four main categories of interaction were identified: 1) object manipulations; 2) content manipulation 3) view manipulation; and 4) information presentation. Specifically, object manipulation requires physical functions and actions such as pointing, moving, selecting objects within the interface, with the aid of an input device such as a mouse or stylus. Two critical components of object manipulation with a pointing device are selection and positioning (Foley, Wallace and Chan 1984) for the successful activation of icons and other screen elements. A selection task involves the user choosing from a set of items on the display (through visual search). A position task consists of the user choosing a single point in a one to three dimensional space (see Jacko, Rosa et al. 1999). These interactions are anticipated to continue to permeate future graphical interfaces (Jacko, Emery et al. 2004).

The Drag & Drop

One instantiation of pointing and selecting indigenous to GUIs and direct manipulation is the aptly named ‘drag-and-drop’ object manipulation. The ‘drag-and-drop’ task has become one of the most prevalent user actions when interacting with information technologies today (Jacko, Moloney et al. 2005). This GUI-based object manipulation adds another level of complexity beyond simple pointing, and selecting (as reviewed by Jacko, Rosa et al. 1999). Drag and drop is characterized by the selection of an object (commonly an icon, such as a ‘file’) and moving it (the ‘drag’) and positioning it on top of (or ‘into’) another object (another icon, such as a ‘folder’) where it is released (the ‘drop’). In terms of functional computing, the drag and drop typically enables the user to change the location of a file to different directories or areas of the file hierarchy. Visual cues within a GUI environment, such as recognizing that the icon is in the correct position to be dropped, and the disappearance of the icon once it is dropped into another, provide feedback to the user during this interaction.

The drag and drop has endured as a common interaction technique for over a decade, even after a significant amount of critical review by the HCI community (see Jacko, Emery et al. 2004 for a discussion). The fundamental nature of the drag and drop manipulation is problematic. Gaver (1989) identified this and coined the term ‘chasing the trash can’ in relation to the drag and drop. The drag and drop is particularly difficult because the user obstructs the target of the drop when they move the icon into place, and it is difficult to ascertain if the object to be dropped is correctly positioned. Even so, the drag and drop is championed as a superior interaction technique to point-click interfaces because of its effectiveness under conditions of high complexity (Joiner, Messer, Light and Littleton 1998). While point and click interactions have proven to be less error prone and more efficient in some instances (Gillan, Hoden, Adam, Rudisill and Magee 1990; MacKenzie, Sellen and Buxton 1991; Joiner, Messer et al. 1998; Smith, Sharit and Czaja 1999), the ever-increasing complexity of GUIS places emphasis on the drag and drop. The following list highlights the most prevalent research areas concerning the drag and drop. While these entries refer to a variety of contexts/applications, they all focus on the drag and drop interaction. Consistent are the keys to a successful drag and drop, including: 1) the visibility and recognition of dynamic icon positions; and 2) the control of the peripheral input device (e.g., the mouse).

Drag and drop research foci:

- The use of auditory feedback (auditory icons) to facilitate more efficient drag and drop (Gaver 1989)
- The application of Fitt's law to drag and drop (Gillan, Hoden et al. 1990).
- A comparison of input devices for both pointing and drag and drop (MacKenzie, Sellen et al. 1991).
- A comparison of pointing and click versus drag and drop for children (Joiner, Messer et al. 1998).
- The use of auditory feedback (earcons) to facilitate drag and drop in both simple and complex task environment (Brewster 1998).
- The impact of aging on the control of input peripherals (Smith, Sharit et al. 1999).
- Drag and drop used as the interaction device for the direct annotation of digital photographs (Shneiderman and Kang 2000).
- A comparison of pointing and click versus drag and drop for children (Inkpen 2001).
- The development of accuracy measures to evaluate pointing devices for interactions such as the drag and drop (MacKenzie, Kauppinen and Silfverberg 2001).
- The evaluation of the effectiveness of multimodal feedback for older adults in a drag and drop (Emery, Edwards et al. 2003).

- The evaluation of the effectiveness of multimodal feedback for individuals with AMD (Jacko, Moloney et al. 2005).

Icons

Icons are the graphical tools and controls, which a user manipulates to execute different computational programs, without inputting any written syntax. “Icons and symbols have been part of the user’s experience of computing for decades, and many people tend to take them for granted as part of the graphical user interface,” (Marcus 2003, p. 37). Icons typically have two components, a pictorial graphic, and a label, usually composed of the file name (Byrne 1993). Several recommendations have been published on the effective design of icons and their use in GUIs. The primary focus of the existing research has been the implication of icon design and use of visual search and icon identification. Identifying characteristics of icons include size, shape, simplicity, quality, and density of icons on the display (including spacing, set size). Sears, Jacko and Robelo (1998) introduce a framework for icon design and provide a useful overview of icon design.

Despite the fact that user interface environments can use 50-100 icons or more in a single application, little work has focused on the characteristic features of icons and the manipulation of these icons as objects in the GUI (The Macintosh Computer, when released in the 1980’s, presented users a corporate suite utilizing approximately 250 icons). Marcus (2003) has asserted that the use of icons will become increasingly critical to the user interface in upcoming years.

Guidelines and loosely formulated standards of icon design are available for developers and designers. The Microsoft Developers Network (MSDN) provides an electronic index that instructs developers on the recommended sizes and colors of icons (Microsoft Developers Network 2005). The website instructs the developers on the importance of effective icon design to communicate their purpose as software objects and that icons should be designed as a set, in relation to each other and the task. According to the MSDN website, icons should be applied in three sizes, 16 x 16 pixels, 32 x 32 pixels, and 48 x 48 pixels, as shown in Figure 1.6.

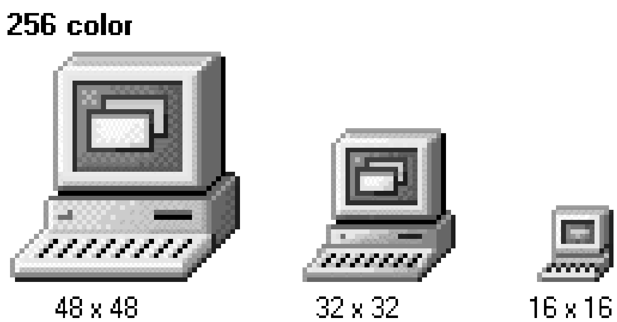


Figure 1.6. Icon size recommendations sample (zoomed) from the MSDN website (Microsoft Developers Network 2005).

Jacko and colleagues, in their review of visual search and selection of icons for users with visual impairments, assessed icon sizes of 9.2 mm, 14.6 mm, 23.2 mm, 36.8 mm and 58.3 icons (based on the relative sizes of letters on a visual acuity eye chart (Jacko, Rosa et al. 1999). At the 9.2 mm size, users with visual impairments were 10 times slower than those individuals without impairments, and for the largest size (58.33 mm) those with visual impairments were just 1.7 times slower. The authors however warn of diminishing returns with an icon size of 58.33 mm, because only six can fit on a 21” monitor without occlusion.

For the task of document identification with visual icons, Byrne asserts that simplicity is critical (Byrne 1993). According to the author, those factors that affect the visual search of icons include general factors, visual factors, and semantic factors. The visual factors include size, color, form, primary icon dimension, spatial organization, and number of objects. For the most part, the systematic investigation of icon design on the user experience is largely unexplored (Byrne 1993), even though the design of icons has been shown to impede task efficiency.

In his experiment, Byrne considered set size, icon type (blank, simple, and complex), and user knowledge about the icon shape and name (through user prompting). The participants in this study were all college students (and visually healthy). Intuitively, Byrne observed that as set size increased, the mean search time also increased. However, the author also observed a significant difference in search time between simple and complex icons in the two largest set sizes (18 and 24). Further work supporting these initial findings (Fleetwood and Byrne 2002, 2003) also revealed that changes in icon type affected the visual strategy employed by the user; high quality icons supported the identification of “clusters” of icons, which users could systematically search. In a more recent paper, Everett and Byrne (2004) extended the results of Byrne’s 1993 work and discovered that in addition to set size and icon simplicity, the effective spacing between icons can induce users to utilize different strategies in visual search. The authors considered screen density, the proportion of the user’s display that contains information. There is a catch-22 when considering icon spacing: while smaller inter-icon spacing can decrease visual encoding search times the increase in density was shown to increase search time. In addition, decreasing density and spacing, by spreading out visual targets, at a certain point creates an inefficient search, because less of the display is viewable in the users' central vision.

In their experiments, Everett and Byrne manipulated icon quality (good, fair and poor), spacing (small = 32 pixels; medium = 64 pixels; large = 96 pixels), and set size (6, 12, 18 and 24). The results of set size and quality were replicated from previous work. Additionally, the effects on search time were observed due to interactions of icon quality and spacing. The authors concluded that spacing and quality can induce users to employ suboptimal search strategies.

Hornoff (2001) has also considered screen density and visual search for a participant population of college students (visually healthy). He concluded that users exhibit a speed accuracy trade off when the space between targets varies. The participants in his study used slower, more accurate strategies as the spacing between icons decreased. The presence of other objects near the target affects search and selection of the object.

Similar to the work by Byrne and Hornoff, Lindberg and Näsänen (2003) consider the impact of both icon size and spacing on visual search and selection. The authors evaluated the effect of changing icon size and spacing, and set size with both visual search times and eye tracking metrics, such as fixation duration. Unfortunately, this study was not rigorously planned or executed: Just three participants were observed; causing the relative power was extremely low. In addition, the conclusions on generalized recommendations for screen design are suspect.

Key Findings

By and large, the results of these studies, however, are relevant to the visual search component of the task only (Byrne 1993; Moyes 1994; Jacko, Rosa et al. 1999; Wiedenbeck 1999; Hornof 2001; Fleetwood and Byrne 2002, 2003; Nasanen and Ojanpaa 2003; Everett and Byrne 2004). The overall objective of the majority of this work is to inform the iterative development of computational models of user interactions, such as ACT-RPM. The investigations and development have yet to execute an in-depth analysis of icon manipulations with peripherals (e.g., drag and drop) in combination with visual search. This is an artifact of the assumption in these models of user skill level and absence of sensory impairments, and often error-free performance. The effect of key icon features such as size, spacing and set size for a drag and drop task is the focus of the experimental task, building on the knowledge of these manipulations on the icon search and selection components of the task.

1.4. Research Questions & Hypotheses

This dissertation is driven by the following, overarching research question: Which characteristics of ocular health and function impact performance in complex interactions with icons (e.g., visual search and drag and drop task in the presence of distracters), to what extent, and under which conditions are the negative effects of AMD best mitigated?

It is the intention this dissertation to examine the following aspects of GUI interaction for individuals with visual impairments attributed to AMD and a set of age-matched, visually healthy controls:

1) Drag and drop in a complex task environment (multiple icons and multiple target distracters), considering the visual search and manipulation of icons

2) Visual augmentations to the interface, considering features of screen density.

    i. Icon size

    ii. Set size

    iii. Spacing

3) Non-visual augmentations to the interface, through the presence and absence of auditory feedback

4) The effect of context through handheld PC vs. desktop PC interactions

5) Each factor considered is assessed via :

    a. Performance measures

        i. Time

        ii. Errors

    b. Physiological measures

        i. Eye movements

        ii. Pupillary response to measure mental workload

    c. Measures of visual function and health

        i. Acuity

        ii. Contrast sensitivity

        iii. Color perception

iv. Visual fields

v. Severity of disease

vi. Visual attention

vii. Perceived visual functioning

d. User interface preferences

The impetus of an investigation of this nature is to contribute substantially to a framework of interaction thresholds for individuals with visual impairments from which universally accessible design guidelines can be derived. To accomplish these improvements to such a framework, several specific hypotheses are formulated. Central to each hypothesis is the quantifiable impact of visual impairments attributed to AMD. These hypotheses are further specified following the introduction of the experiment and variables at the conclusion of Chapter 2.

Hypothesis 1

For all users, there is a point of diminishing return for performance gains attributed to increases in icon size, set size, spacing, and overall screen density.

Hypothesis 2

The potential positive influence of auditory feedback on the drag and drop task is effectively masked by the complexity of the task (multiple icons and multiple targets as compared to the single file – single folder task used in previous studies).

Hypothesis 3

Measures of ocular health and visual function are predictors of performance in the required task.

Hypothesis 4

The interactions of those users in the handheld PC group will be less efficient than those users in the desktop PC group.

1.5 REFERENCES

Ahn, S. J. and D. K. Legge (1995). "Psychophysics of reading-XII: Predictors of magnifier-aided reading speed in low vision." Vision Research **35**: 1935-1938.

Alberti, W. E., G. Richard and R. H. Sagerman (2000). Age-Related Macular Degeneration: Current Treatment and Concepts. New York, NY, Springer-Verlag.

Amar, R., S. Dow, R. Gordan, M. R. Hamid and C. Sellers (2003). Mobile ADVICE: An accessible device for visually impaired capabiltiy enhancement. CHI 2003, Ft. Lauderdale, FL, ACM.

American Foundation for the Blind. (2004). "Statement on technology and aging: statement for the record of the American Foundation for the Blind prepared for the United States Senate Special Committee on Aging regarding assistive technologies for independent living: Opportunities and challenges, April 27, 2004 hearing." Retrieved July 8, 2004, from http://www.afb.org.

American Macular Degeneration Foundation. (2002). "What is macular degeneration?" Retrieved November 14, 2002, from http://www.macular.org/disease.html.

Arditi, A. (2002). "Web Accessibility and Low Vision." Aging & Vision **14**(2): 2-3.

Arditi, A. (2003). "Low vision web browsing and allocation of screen space resources." Investigative Ophthalmology and Visual Science **44**(5): 2767.

Arditi, A. (2004). "Adjustable typography: an approach to enhancing low vision text accessibility." Ergonomics **47**(5): 469-482.

Bailey, I. L. and M. Bullimore (1991). "A new test of disability and glare." Optometry and Visual Science **68**: 911-917.

Bailey, I. L. and A. Hall (1989). Visual Impairment: An Overview. New York, NY, American Foundtion for the Blind.

Baker, P. M. A. (2004). "Factors influencing adoption of wireless technologies: Key issues, barriers and opportunities for people with disabilities." Retrieved July 20, 2004, from http://www.wirelessrerc.gatech.edu/news/policyassessment.html.

Barnard, L., J. S. Yi, J. A. Jacko and A. Sears (2004). "Capturing the effects of context on human performance in mobile computing systems." Submitted for publication in Personal and Ubiquitous Computing.

Barnard, L., J. S. Yi, J. A. Jacko and A. Sears (2004). "An empirical comparison of use-in-motion evaluation scenarios for mobile computing devices." Submitted for publication in the International Journal of Human Computer Studies.

Becker, S. and D. Lundman (1998). Improving access to computers for blind and visually impaired persons: The development of a test method for usability. Telemeatics for the Integration of Disabled and Elderly People (TIDE) 1998 Conference, Helsinki.

Ben-Shakhar, G. a. S., Limor (2001). "The relationship between the ability to divide attention and standard measures of general cognitive abilities." Intelligence **29**(2001): 294-306.

Bergstein, B. (2003). Better living through technology. Associated Press. New York.

Biglan, A. W., V. B. Van Hasselt and J. Simon (1988). Visual Impairment. Handbook of Developmental and Physical Disabilities. V. B. Van Hassek. New York, NY, Pergamon Press**:** 471-502.

Brewster, S. (1998). Sonically-enhanced drag and drop. ICAD 1998, Glasgow.

Brewster, S., P. C. Wright and A. D. N. Edwards (1994). The design and evaluation of an auditory-enhanced scrollbar. 1994 ACM Conference on Human Factors in Computing Systems (CHI'94), Boston, MA, ACM Press.

Brinker, D. and P. J. Beek (1996). "Reading with magnifiers." Ergonomics **39**(1231-1248).

Byrne, M. D. (1993). Using icons to find documents: Simplicity is critical. Proceedings of the SIGCHI conference on Human factors in computing systems, Amsterdam, The Netherlands.

Center on Aging Society. (2002). "Visual Impairments
A growing concern as population ages." Georgetown University Data File: Chronic and Disabling Conditions fo rhte 21st Century 3. Retrieved August 1, 2004, 2003, from http://ihcrp.georgetown.edu/agingsociety/pdfs/visual.pdf.

Coeckelbergh, T. R. M., F. W. Cornelissen, W. H. Brouwer and A. C. Kooijam (2002). "The effect of visual field defects on eye movement and practical fitness to drive." Vision Research **42**: 669-677.

Collin, H. B., C. Han and P. C. Korh (1988). "Age changes in the visual field using the Humphrey Field Analyzer." Clinical and Experiemental Optometry **71**: 174-178.

Connell, B. R., M. Jones, R. Mace, J. Mueller, A. Mullick, E. Ostroff, J. Sanford, E. Steinfeld, M. Story and G. Vanderheiden. (1997). "The principles of universal design." Retrieved July 29, 2004, from http://www.design.ncsu.edu:8120/cud/univ_design/principles/udprinciples.htm.

Craven, J. (2003). "Access to electronic resources by visually impaired people." Information Research **8**(4).

Dix, A., J. Finlay, G. Abowd and R. Beale (1998). Human-computer interaction. New York, NY, New York, NY: Prentice Hall.

Dunlop, M. and S. Brewster (2002). "The challenge of mobile devices for human computer interaction." Personal and Ubiquitous Computing **6**: 235-236.

Eason, K. D., C. Johnson and S. Fairclough (1991). The interaction of sensory modalities in the man-machine interface. Vision and dysfunction: The man-machine interface. J. A. J. R. (Ed.). Boca Raton, FL, CRC.

Edwards, P. J., L. Barnard, V. K. Emery, J. S. Yi, K. P. Moloney, T. Kongnakorn, J. A. Jacko, F. Sainfort, P. R. Oliver, J. Pizzimenti, A. Bade, G. Fecho and J. Shallo-Hoffmann (2004). Strategic Design for Users with Diabetic Retinopathy:

Factors Influencing Performance in a Menu-Selection Task. The Sixth International ACM SIGCAPH Conference on Assistive Technologies (ASSETS 2004), Atlanta.

Edwards, P. J., L. Barnard, V. K. Emery, J. S. Yi, K. P. Moloney, T. Kongnakorn, J. A. Jacko, F. Sainfort, P. R. Oliver, J. Pizzimenti, A. Bade, G. Fecho and J. Shallo-Hoffmann (2004). Strategic Design for Users with Diabetic Retinopathy:Factors Influencing Performance in a Menu-Selection Task. Sixth International ACM SIGCAPH Conference on Assistive Technologies (ASSETS 2004), Atlanta.

Edwards, P. J., L. Barnard, V. K. Leonard, J. S. Yi, K. P. Moloney, T. Kongnakorn, J. A. Jacko and F. Sainfort (2005). "Understanding users with Diabetic Retinopathy: Factors that affect performance in a menu selection task." Behaviour & Information Technology **24**(3): 175-186.

Emery, V. K., P. J. Edwards, J. A. Jacko, K. P. Moloney, L. Barnard, T. Kongnakorn, F. Sainfort and I. U. Scott (2003). Towards achieving universal usability for older adults through multimodal feedback. ACM Conference on Universal Usability (CUU), Vancouver, British Columbia, Canada.

Emery, V. K., J. A. Jacko, T. Kongnakorn, V. Kuruchittham, S. Landry, G. M. N. III, A. Sears and J. Whittle (2001). Identifying critical interaction scenarios for innovative user modeling. 10th International Conference on Human-Computer Interaction, New Orleans, LA.

Emery, V. K., J. A. Jacko, F. Sainfort and J. S. Yi (in press). Human Factors and Ergonomic Methods. Handbook for Human Factors and Ergonomics. G. Salvendy. New York, Taylor & Francis.

Everett, S. P. and M. D. Byrne (2004). "Unintented effects: Varying icon spacing changes users' visual search strategy." CHI Letters **6**(1): 695-702.

EyeMDLink.com. (2002). "Photodynamic Therapy for Macular Degeneration." Retrieved July 30, 2004, from http://www.eyemdlink.com/NewsArticle.asp?NewsID=39.

Farnsworth, D. (1947). The Farnsworth Dichotomous Test for Color Blindness: Manual. New York, The Psychological Corporation.

Farrell, F. E. (1991). Fitting physical screen parameters to the human eye. Vision and visual dysfunction: The man-machine interface. J. A. J. Roufs. Boca Raton, FL, CRC.

Ferris, F. L., III, A. Kassoff and G. H. Bresnick (1982). "New visual acuity charts for clinical research." American Journal of Ophthalmology **94**: 91-96.

Fleetwood, M. D. and M. D. Byrne (2002). "Modeling icon search in ACT-R/PM." Cognitive Systems REsearch **3**: 15-33.

Fleetwood, M. D. and M. D. Byrne (2003). Modeling visual search of displays: A revised ACT-R/PM model of icon search based on eye tracking and experimental data. The Fifth International Conference on Cognitive Modeling.

Foley, J. D., V. L. Wallace and P. Chan (1984). "The human factors of computer graphics interaction techniques." IEEE Computer Graphics and Applications **89**(8): 13-48.

Fortuin, F. T. and S. Omata. (2004). "Designing universal user interfaces - The application of universal design rules to eliminate information barriers for the visually impaired and the elderly." Retrieved May 19, 2004, from http://www.visionconnection.org.

Fraser, J. and C. Gutwin (2000). A framework of assistive pointers for low vision users. ACM Conference on Assistive Technologies, Arlington, VA.

Gaver, W. (1989). "The SonicFinder: An interface that uses auditory icons." Human Computer Interaction **4**(1): 67-94.

Gerber, E. and C. Kirchner (2001). "Who's surfing? Internet access and computer use by visually impaired youths and adults." Journal of Visual Impairment & Blindness **2001**(March): 176-181.

Gillan, D. J., K. Hoden, S. Adam, M. Rudisill and L. Magee (1990). How does Fitt's law fit pointing and dragging? SIGCHI Conference on Human Factors in Computing Systems: Empowering People, New York, ACM Press.

Harper, S., C. Goble and R. Stevens (2001). "Web mobility guidelines for visually impaired surfers." Journal of Research and Practice in Information Technology **33**(1): 30-41.

Hornof, A. (2001). "Visual search and mouse-pointing in two-dimensional visual hierarchies." ACM Transactions on Computer-Human Interaction **8**(3): 171-197.

Inkpen, K. M. (2001). "Drag-and-drop versus point-and-click mouse interaction styles for children." ACM Transactions on Computer-Human Interaction **8**(1): 1-33.

Jacko, J. A. (1999). "Incorporating clinical diagnoses in the prediction of performance on computer-based tasks for low vision users." ACM SIGCAPH Computers and the Physically Handicapped **64**(June 1999): 3-5.

Jacko, J. A. (2000). "Visual impairment: The use of visual profiles in evaluations of icon use in computer-based tasks." International Journal of Human-Computer Interaction **12**(1): 151-164.

Jacko, J. A. (2000). "Visual impairment: The use of visual profiles in evaluations of icon use in computer-based tasks." International Journal of Human-Computer Interaction. **12**(1): 151-164.

Jacko, J. A., L. Barnard, T. Kongnakorn, K. P. Moloney, P. J. Edwards, V. K. Emery and F. Sainfort (2004). "Isolating the effects of visual impairment: Exploring the effect of AMD on the utility of multimodal feedback." CHI Letters **6**(1): 311-318.

Jacko, J. A., L. Barnard, J. S. Yi, P. J. Edwards, V. K. Leonard, T. Kongnakorn, K. P. Moloney and F. Sainfort (in press). "Empirical validation of the Windows accessibility settings and multimodal feedback for a menu selection task for users with diabetic retinopathy." Behaviour & Information Technology.

Jacko, J. A., A. B. Barreto, J. Y. M. Chu, H. S. Bautsch, G. J. Marmet, I. U. Scott and R. H. Rosa (2000). Using eye tracking to investigate graphical elements for fully sighted and low vision users. ACM conference on Eye Tracking Research Applications (ETRA), Palm Beach, FL.

Jacko, J. A., A. B. Barreto, G. J. Marmet, J. Y. M. Chu, H. S. Bautsch, I. U. Scott and R. H. Rosa (2000). Low vision: The role of visual acuity in efficiency of cursor movement. ACM Conference on Assistive Technologies, Arlington, Virginia.

Jacko, J. A., A. B. Barreto, I. U. Scott, J. Y. M. Chu, H. S. Vitense, F. T. Conway and W. B. Fain (2002). "Macular degeneration and visual icon use: Deriving guidelines for improved access." Universal Access and the Information Society **1**: 197-296.

Jacko, J. A., V. K. Emery, P. J. Edwards, M. Ashok, L. Barnard, T. Kongnakorn, K. P. Moloney, F. Sainfort and I. U. Scott (2004). "The effects of multimodal feedback on older adults' task performance given varying levels of computer experience." Behaviour & Information Technology **23**(4): 247-264.

Jacko, J. A., K. P. Moloney, T. Kongnakorn, L. Barnard, P. J. Edwards, V. K. Leonard, F. Sainfort and I. U. Scott (2005). "Multimodal feedback as a solution to ocular disease-based user performance decrements in the absence of functional visual loss." International Journal of Human-Computer Interaction **18**(1): 183-218.

Jacko, J. A., R. H. Rosa, I. U. Scott, C. J. Pappas and M. A. Dixon (1999). Linking visual capabilities of partially sighted computer users to psychomotor task performance. 9th International Conference on Human-Computer Interaction, Munich, Germany.

Jacko, J. A., I. U. Scott, A. B. Barreto, H. S. Bautsch, J. Y. M. Chu and W. B. Fain (2001). Iconic visual search strategies: A comparison of computer users with AMD versus computer users with normal vision. 10th International Conference on Human-Computer Interaction, New Orleans, LA.

Jacko, J. A., I. U. Scott, F. Sainfort, L. Barnard, P. J. Edwards, V. K. Emery, T. Kongnakorn, K. P. Moloney and B. Zorich (2003). "Older adults and visual impairment: What do exposure times and accuracy tell us about performance gains associated with multimodal feedback?" CHI Letters **5**(1): 33-40.

Jacko, J. A., I. U. Scott, F. Sainfort, L. Barnard, P. J. Edwards, V. K. Emery, T. Kongnakorn, K. P. Moloney and B. S. Zorich (2003). Older adults and visual impairment: What do exposure times and accuracy tell us about performance gains associated with multimodal feedback? ACM Conference on Human Factors in Computing Systems (CHI 2003), Ft. Lauderdale, FL.

Jacko, J. A., I. U. Scott, F. Sainfort, K. P. Moloney, T. Kongnakorn, B. Zorich and V. K. Emery (2002). Effects of multimodal feedback on the performance of older adults with normal and impaired vision. Lecture notes in computer science. C. S. Noelle Carbonell. NEED CITY, Springer. **2615:** 3-22.

Jacko, J. A. and A. Sears (1998). Designing interfaces for an overlooked user group: Considering the visual profiles of partially sighted users. ACM Conference on Assistive Technologies, Marina del Rey, California.

Jacko, J. A. and H. S. Vitense (2001). The foundation of a conceptual framework for individuals with disabilities. 10th International Conference on Human-Computer Interaction, New Orleans, LA.

Joiner, R., D. Messer, P. Light and K. Littleton (1998). "It is best to point for young children: A comparison of children's pointing and dragging." Computers in Human Behavior **14**(3): 513-529.

Kaufman, P. L. and A. Alm (2003). Adler's Physiology of the Eye: Clinical Application. St. Louis, MO, Mosby.

Kline, D. W. and C. T. Scialfa (1997). Sensory and perceptual functioning: Basic research and human factors implications. Handbook of Human Factors and the Older Adults. A. D. Fisk and W. A. Rodgers. New York, Academic Press**:** 28-54.

Kline, R. L. and E. P. Glinert (1995). Improving GUI accessibility for people with low vision. SIGCHI Conference on Human Factors in Computing Systems, Denver, Colorado.

Kornblum, J. (2004). Studies foresee increased vision loss among boomers; Proper treatment could lower cases. USA Today**:** D.07.

Lee, H.-W., G. E. Legge and A. Oritz (2003). "Is word recognition different in central and peripheral vision?" Vision Research **43**: 2837-2846.

Lin, R. T. (2004). "Computer users find relief with lubricating eye drop." Ophthalmology Times **2004**(March 15).

Lindberg, T. and R. Nasanen (2003). "The effect of icon spacing and size on the speed of icon processing in the human visual system." Displays **24**(2003): 111-120.

Ludi, S. (2000). Animated icons: Re-inventing visual cues for the visually impaired computer user. ACM Conference on Universal Usability, Arlington, VA.

MacKenzie, I. S., T. Kauppinen and M. Silfverberg (2001). "Accuracy measures for evaluating computer pointing devices." CHI Letters **v3**: 9-16.

MacKenzie, I. S., A. Sellen and W. A. S. Buxton (1991). A comparison of input devices in elemental pointing and dragging tasks. ACM Conference on Human Factors in Computing Systems: Reaching Through, New York, ACM Press.

Macular Degeneration Partnership. "Anatomy of the Eye." Retrieved May 5, 2003, from http://amd.convio.net/site/PageServer?pagename=Anatomy_of_the_Eye.

Macular Degeneration Partnership. "What is AMD?" Retrieved November 5, 2002, from http://www.macd.net/amd/index.asp.

Marcus, A. (2003). Icons, symbols and signs: Visible languages to facilitate communication. Interactions**:** 37-43.

McNeil, A. and L. Griffin (2002). Policy and regulatory assessment, Factors influencing adoption of wireless technologies: key issues, barriers and opportunities for people with disabilities.

MD Foundation. (2004). "Pictorials Descriptions of AMD." Retrieved July 29, 2004, from http://www.eyesight.org/Pictorials/Pic-Look_Like/pic-look_like.html.

Microsoft Developers Network. (2005). "Icon Guidelines." from www.msdn.com.

Morrell, R. W. (2002). Forward. Older adults, health information, and the world wide web. R. W. Morrell. Mahwah, NJ, Earlbaum.

Moyes, J. (1994). When users do and don't rely on icon shape. Proceedings of the SIGCHI conference on Human factors in computing systems, Boston, MA.

Nasanen, R. and H. Ojanpaa (2003). "Effect of image contrast and sharpness on visual search for computer icons." Displays **24**: 137-144.

National Eye Institute. (2001). "Are you at risk for Age-Related Macular Degeneration?" Retrieved November 4, 2002, 2002, from http://www.nei.nih.gov/health/maculardegen/armd_risk.htm.

Nelson-Quigg, J. M., K. Cello and C. A. Johnson (2000). "Predicting binocular visual field sensitivity form monocular visual field results." Investigative Ophthalmology and Visual Science **41**(8): 2212-2221.

Newburger, E. C. (2001). Home computers and internet use in the United States. USA, US Census Bureau.

Newell, A. F. and P. McGregor (1997). Human-computer interfaces for people with disabilities. Handbook of human-computer interaction. M. G. Helander, T. K. Landauer and P. V. Prabhu. Amsterdam, The Netherlands, Elsevier**:** 813-824.

Oakley, I., S. A. Brewster and P. D. Gray (2001). Solving multi-target haptic problems in menu interaction. CHI '01 Extended Abstracts on Human Factors in Computing Systems, Seattle, Washington, ACM.

Orr, A. L. (1998). Issues in Aging and Vision. New York, American Foundation for the Blind.

Owsley, C. and M. E. Sloan (1990). Vision and aging. Handbook of neuropsychology. F. Boller and J. Grafman. Amsterdam, Elsevier. **4:** 229-249.

Pelli, D. G., J. G. Robson and A. J. Wilkins (1988). "The design of a new letter chart for measuring contrast sensitivity." Clinical Visual Science **2**: 187-199.

Quillen, D. (1999). "Common causes of vision loss in elderly patients." American Family Physician **60**(1): 99-108.

Sanders, M. S. and E. J. McCormick (1993). Human Factors in Engineering and Design. New York, McGraw-Hill.

Schieber, F. (1994). Recent developments in vision, aging and driving: 1988-1994, Transportation Research Board.

Scott, I. U., W. J. Feuer and J. A. Jacko (2002). "Impact of graphic user interface screen features on computer task accuracy and speed in a cohort of patients with Age-related Macular Degeneration." American Journal of Ophthalmology **134**(6): 857-862.

Scott, I. U., W. J. Feuer and J. A. Jacko (2002). "Impact of visual function on computer task accuracy and reaction time in a cohort of patients with Age-related Macular Degeneration." American Journal of Ophthalmology **133**(3): 850-857.

Sears, A., J. A. Jacko, B. Brewer and L. Robelo (1998). A framework for visual icon design. 42nd Annual Meeting of the Human Facotrs Societ, Santa Monica, CA, Human Factors Society.

Sears, A., M. Lin, J. Jacko and Y. Xiao (2003). When Computers Fade.Pervasive Computing and Situationally-Induced Impairments and Disabilities. Human-Computer Interaction: Theory and Practice (Part II). C. Stephanidis and J. Jacko, Lawrence Erlbaum Associates**:** 1298-1302.

Shneiderman, B. (1998). Designing the user interface: Strategies for effective human-computer interaction. Reading, MA, Addison-Wesley.

Shneiderman, B. and H. Kang (2000). "Direct annotation: A drag-and-drop strategy for labeling photos." International Conference on Information Visualisation (IV2000): 88-98.

Silva, L., O. Regina and P. Bellon (2002). A novel approach to aid low vision computer users. K. Miesenberger, J. Klaus and W. Zagler. Berlin, Heidelberg. **2398**.

Smith-Jackson, T. L., M. A. Nussabaum and A. M. Mooney (2003). "Accessible cell phone design: development and application of a needs analysis framework." Disability and Rehabilitation **25**(10): 549-560.

Smith, M. W., J. Sharit and S. J. Czaja (1999). "Aging, motor control and the performance of computer related mouse tasks." Human Factors **41**(3): 389-396.

Tun, P. A. and A. Wingfield (1997). Language and communication: funndmentals of speech communication and language processing in old age. Handbook of Human Factors and the Older Adult. A. D. Fisk and W. A. Rodgers. New York, Academic Press**:** 125-149.

Vanderheiden, G. C. (2004). "Use of audio-haptic interface techniques to allow nonvisual access to touchscreen appliances." Retrieved July 20, 2004, from http://trace.wisc.edu/docs/touchscreen/chi_conf.htm.

VisionChannel.net. (no date). "VisionChannel.net. (n. d.). Macular Degeneration: Treatment." Retrieved July 29, 2003.

Vitense, H. S., J. A. Jacko and V. K. Emery (2002). "Foundation for improved interaction by individuals with visual impairments through multimodal feedback." Universal Access and the Information Society **2002**(2): 76-87.

Vitense, H. S., J. A. Jacko and V. K. Emery (2003). "Multimodal feedback: an assessment of performance and mental workload." Ergonomics **46**(1-3): 68-87.

Whittaker, S. G. (1998). Choosing assistive devices when computer users have impaired vision. CSUN.

Wiedenbeck, S. (1999). "The use of icons and labels in an end user application program: an empirical study of learning and retention." Behaviour & Information Technology **18**(2): 68-82.

CHAPTER 2

METHODOLOGIES

2.1. Participants

Participants were recruited from the Nova Southeastern University (NSU) College of Optometry patient pool and through friends and colleagues of NSU faculty and staff. After physicians made initial contact with patients by letter, recruitment took place over the phone from Atlanta.

The inclusion criteria were:

1. Presence of AMD (initially based on medical record, but screened the day of testing through ophthalmic exam to verify AMD as their only diagnosis).
2. Absence of other ocular conditions (e.g., cataract, glaucoma, etc.);
3. As a minimum, high school education;
4. English fluent; and
5. Computer experience score indicative of frequent use and/or familiarity with a breadth of applications (derived by Emery, Edwards, Jacko et al. 2003).

In addition to the patients with AMD, age-matched, visually healthy controls were recruited. This cohort was subject to the same inclusion criteria as the AMD group, but was not diagnosed with AMD. In total, 27 participants were recruited for this study and participated in the computer task. Table 2.2 details the participants' demographic backgrounds, including a summary of ages, and the breakdown of the number of participants according to gender, ethnicity, education level, and if they were currently working or retired. Table 2.2 also provides a breakdown of participants and their demographics according to diagnosis (Control group vs. AMD) and according to which interface condition the participant assigned under (desktop vs. handheld).

Non-parametric statistical analyses were applied to determine if differences existed between these different cohorts for any of the demographics. Table 2.1 summarizes these results. The Mann-Whitney non-parametric test confirmed there were no statistically significant differences between the handheld and desktop cohorts for age ($p = .054$). No significant differences in age were observed between the controls and AMD cohorts within either the desktop ($p = .368$) or the handheld ($p = .568$) experimental conditions.

Chi Squared Tests were performed to examine differences in the categorical demographics. The test revealed no significant differences between the handheld and desktop cohorts, nor between the Control and AMD groups within each condition on Gender, Race, or Working History (for alpha = .10). These results ensure that any differences that emerge in the data analyses are not due to those factors listed, and enable a more valid comparison of interaction between the platforms and groups.

Table 2.1. Statistical Comparisons of demographics between cohorts.

Demographic	Statistical Test	Comparison		
		Desktop vs. Handheld	*Desktop: Control vs. AMD*	*Handheld: Control vs. AMD*
Age	Mann-Whitney	Mann-Whitney U (N = 27)= 51, Z = -1.94, p =.054	Mann-Whitney U (N = 14)= 5.5 Z = -1.74, p =.088	Mann-Whitney U (N = 13) = 13.0, Z = -.772, p =.503
Gender	Chi-Squared	X^2 (2, N= 27) = 1.84, p = .175	X^2 (2, N= 14) = .117 p = .733	X^2 (2, N= 13) < .001, p = 1.00
Race		X^2 (2, N= 27) = .005, p = .946	X^2 (2, N= 14) = 1.12, p = .290	X^2 (2, N= 13) = .037, p = .848
Education Level		X^2 (3, N= 27) = 1.54, p = .464	X^2 (3, N= 14) = .661, p = .719	X^2 (3, N= 13) = .481, p = .786
Working History		X^2 (2, N= 27) = .909, p = .340	X^2 (2, N= 14) = .018, p = .894	X^2 (2, N= 13) = .174, p = .676

Computer experience served as a key inclusion criterion, as it has been shown in previous work to have significant impact older adults interactions with information technology and efficacy of design interventions (Emery, Edwards et al. 2003; Jacko, Emery, Edwards et al. 2004). Table 2.3 summarizes the computer experience of the participants, according to desktop and handheld conditions, control and AMD cohorts. In their screening, participants were asked to rate their comfort with computers as v*ery comfortable, comfortable, neither comfortable nor uncomfortable,* or *very uncomfortable.*

Table 2.2. Participant Demographic Summary: Summary statistics are provided for Age, and incidence of participants per category provided for Gender, Race, Education and Working History.

	N	**Age** (years)	**Gender**	**Race**	**Education Level**	**Working History**
All Participants	**27**	Age Range: 53-90 Mean Age (standard error): 72.7 (1.83) Median Age: 75.0	Male: 13 Female: 14	Caucasian: 24 Black: 3	High School: 7 College: 14 Post-graduate: 6	Currently Working: 9 Retired: 18
Desktop	***14***	Age Range: 62-90 Mean Age (standard error): 76.5 (2.08) Median Age: 77	Male: 9 Female: 5	Caucasian: 13 Black: 1	High School: 5 College: 6 Post-graduate: 3	Currently Working: 2 Retired: 11
Controls	*2*	Age Range: 62-72 Mean Age (standard error): 67.00 (5.00) Median Age: 67	Male: 2 Female: 0	Caucasian: 1 Black: 1	High School: 1 College: 1 Post-graduate: 0	Currently Working: 1 Retired: 1
AMD	*12*	Age Range: 67-90 Mean Age (standard error): 77.08 (1.99) Median Age: 77.5	Male: 7 Female: 5	Caucasian: 12 Black: 0	High School: 4 College: 5 Post-graduate: 3	Currently Working: 2 Retired: 10
Handheld	***13***	Age Range: 53-82 Mean Age (standard error): 68.69 (2.74) Median Age: 70	Male: 4 Female: 9	Caucasian: 11 Black: 2	High School: 2 College: 8 Post-graduate: 3	Currently Working: 6 Retired: 7
Controls	*4*	Age Range: 57-75 Mean Age (standard error): 66.75 (3.71) Median Age: 67.5	Male: 1 Female: 3	Caucasian: 4 Black: 0	High School: 1 College: 2 Post-graduate: 1	Currently Working: 1 Retired: 3
AMD	*9*	Age Range: 53-82 Mean Age (standard deviation): 69.56 (3.69) Median Age: 73	Male: 3 Female: 6	Caucasian: 7 Black: 2	High School: 1 College: 6 Post-graduate: 2	Currently Working: 5 Retired: 4

The computer experience score was derived from each participant's frequency of computer use and the number of applications with which they were familiar, consistent with previous studies (Emery, Edwards et al. 2003; Edwards, Barnard, Emery et al. 2004; Jacko, Emery et al. 2004). This method is useful for a population with visual impairments who may have computer experience but have had to abandon use due to increases in visual dysfunction.

In addition, participants were also surveyed on their ownership of handheld and mobile computers. Just one participant reported ownership of a handheld PC, which he had owned for approximately 3 years. This participant reported using the handheld PC for both scheduling and for the address book function. As an alternative means to derive familiarity with small, mobile computers, participants were also asked to report the number of years they had owned a cell phone. These results are also summarized in Table 2.3.

Similar to the examination of demographic data, statistical analyses were performed confirm the absence of significant differences between the cohorts for computer experience, and are summarized in Table 2.4. As with the demographic analyses, there were no significant differences between the cohorts or experimental conditions based on the computer experience variables (X^2, p >.146; Man Whitney U, p > .280). This ensured that the groups were equivalent on computer experience. This reinforces the validity of this investigations' ability to isolate differences in interaction due to ocular function, disease or technology platform and to other naturally variable factors.

Table 2.3. Summary of participant computer experience.

Cohort	N	Computer Comfort Rating	Computer Experience Score	Experience With Cell Phone
All Participants	**27**	Very comfortable: 11 Comfortable: 10 Neither: 3 Uncomfortable: 2 Very uncomfortable: 1	Mean (Standard Error): 8.07 (.751) Median: 9	Never Owned: 11 Owned 2-5 years: 9 Owned 10 years or longer: 7
Desktop	***14***	Very comfortable: 6 Comfortable: 4 Neither: 3 Uncomfortable: 0 Very Uncomfortable: 1	Mean (Standard Error): 7.07 (1.14) Median: 6.5	Never Owned: 6 Owned 2-5 years: 5 Owned 10 years or longer: 3
Controls	*2*	Very comfortable: 2 Comfortable: 0 Neither: 0 Uncomfortable: 0 Very Uncomfortable: 0	Mean (Standard Error): 8.00 (5.00) Median: 8.00	Never Owned: 1 Owned 2-5 years:0 Owned 10 years or longer: 1
AMD	*12*	Very comfortable: 4 Comfortable: 4 Neither: 3 Uncomfortable: 0 Very Uncomfortable: 1	Mean (Standard Error): 6.92 (1.18) Median: 6.5	Never Owned: 5 Owned 2-5 years: 5 Owned 10 years or longer: 2
Handheld	***13***	Very comfortable: 5 Comfortable: 6 Neither: 0 Uncomfortable: 2 Very Uncomfortable: 0	Mean (Standard Error): 9.15 (.912) Median: 10	Never Owned: 5 Owned 2-5 years: 4 Owned 10 years or longer: 4
Controls	*4*	Very comfortable: 0 Comfortable: 4 Neither: 0 Uncomfortable: 0 Very Uncomfortable: 0	Mean (Standard Error): 8.25 (2.17) Median: 9.5	Never Owned: 3 Owned 2-5 years: 0 Owned 10 years or longer: 1
AMD	*9*	Very comfortable: 5 Comfortable: 2 Neither: 0 Uncomfortable: 2 Very Uncomfortable: 0	Mean (Standard Error): 9.56 (.973) Median: 10	Never Owned: 2 Owned 2-5 years: 4 Owned 10 years or longer: 3

Table 2.4. Statistical Comparisons of computer experience between cohorts.

Demographic	Statistical Test	Comparison		
		Desktop vs. Handheld	*Desktop: Control vs. AMD*	*Handheld: Control vs. AMD*
Computer Experience Score	Mann-Whitney	Mann-Whitney U (N = 27)= 68.5, Z = -1.098, p =.280	Mann-Whitney U (N = 14)= 10.0, Z = -.368, p =.713	Mann-Whitney U (N = 13)= 15.5, Z = -.394, p =.710
Computer Comfort*	Chi-Squared	X^2 (3, N= 27) = 6.46, p = .167	X^2 (3, N= 14) = .933 p = .627	X^2 (2, N= 13) = .037 p = .848
Cell Phone Use		X^2 (3, N= 27) = .308, p = .857	X^2 (2, N= 14) = 1.75 p = .417	X^2 (2, N= 13) = 3.85, p = .146

*In order to meet the assumptions for Chi-Squared, the *Computer* Comfort response was recoded, so that the participants who responded with V*ery Comfortable* and C*omfortable* were aggregated to make up a single *Comfortable* group. Likewise, *Very Uncomfortable* ratings were grouped into the *Uncomfortable r*esponses to form a single *Uncomfortable* group. In the handheld condition, none of the participants rated their comfort as *neither.*

2.2. <u>Experimental Protocol</u>

Clinical Methods & Summaries

Aside from recruitment, all testing took place at NOVA Southeastern University (NSU) College of Optometry and Eye Clinic in Fort Lauderdale, Florida. As incentive for their participation, participants were given US $50 and a comprehensive ocular examination by NSU Optometrists and faculty. The examination of the participants included the assimilation of general intake information such as medical history and demographics, in addition to a comprehensive ocular examination that assessed aspects of both ocular health and ocular function. The following tests were executed and information gathered by the Nova team with the participants:

Background/Intake

Demographics

- Age
- Gender
- Race
- Smoking history
- Driving and vehicular accident history

Medical History

- Current medications
- Co morbidities (systemic disease)

Vitals

- Blood pressure
- Heart rate

Ocular health history

- Age of AMD diagnosis
- Family history of AMD
- Ocular surgery

Ophthalmic Exam

Functional Vision Assessment

- Refractive Error for distance and reading (i.e., current prescription of glasses and best corrected prescription)
- ETDRS distance visual acuity assessment*
- ETDRS near visual acuity assessment
- Pelli-Robson contrast sensitivity assessment
- D-15 Color vision test
- American Optical Hardy-Rand-Rittler plates (HRR) Color vision test
- Amsler Grid
- Pupillary light reflex
- Preferential Hyper-acuity Perimeter (PHP)

Ocular Health Assessment/Diagnosis

- Intraocular pressure
- Examination of lens
- Examination of the retina
- Digital photos of the retina (fundas images)
- Retinal thickness analysis

As this study represented a collaborative research effort with the NOVA team, the entire set of clinically-acquired measures of visual health and function collected are not relevant to this dissertation’s objectives. For those measures applicable to the dissertation and hypothesis, a more comprehensive explanation follows. For precise definitions of these ocular functions, please refer to Table 1.2, which is found in Chapter 1 of this dissertation.

Participants’ visual acuity (VA) was evaluated for both distance (20 feet) and near (30 cm) visual function. VA has been noted as one of the most extensively used tests of visual function and to monitor disease progression, such as AMD (Bressler, Bressler, West, Fine and Taylor 1988; Arditi and Cagenello 1993; Bird, Bressler, Bressler et al. 1995).To assess distant VA, NSU optometrists utilized the ETDRS chart and method, from the early treatment of diabetic retinopathy study (International Council of Ophthalmology 2005). ETDRS is regarded as the gold standard for visual acuity assessment of the National Eye Institute and the National Institute of Health. A sample ETDRS acuity chart is provided in Figure 2.1. A score is given based on the line which contains the smallest letters that can be perceived, as well as the number of letters accurately perceived. The scores generated from the ETDRS assessment were converted to Snellen acuity (e.g., 20/20, 20/30, 20/45). The greater the denominator in the Snellen score, the worse the visual acuity. NSU optometrists assessed visual acuity for each eye independently as well as both eyes together (a.k.a. binocular vision).

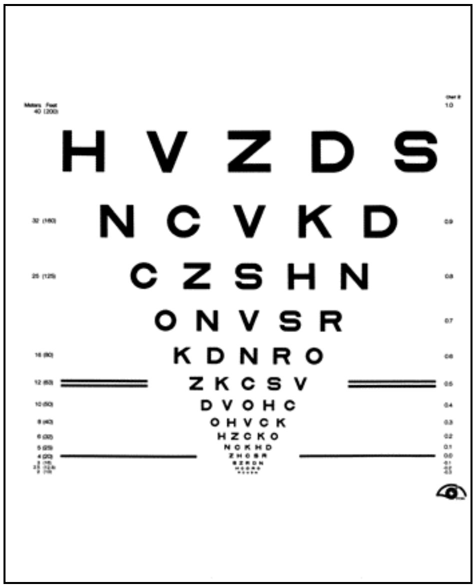


Figure 2.1. Sample ETDRS acuity test (source: National Eye Institute).

Near visual acuity (NVA) was assessed using the Lighthouse Continuous Text card. This card assesses the size of text that could be read at a chosen distance of 30 cm (13 inches), with 4.0 M print that is equivalent to 5.0 mm letters or slightly under 2" tall lower case print (Ormerod and Mussatt 2005). NVA scores were obtained for the left and right eyes independently and also the binocular vision (both eyes). As with distance VA, the NVA scores were also converted to Snellen Acuity Scores. Table 2.5 provides a summary of the visual acuity scores exhibited by the participants in this study, structured according to experimental groups. This table reports the denominator of the Snellen acuity scores. An increasing value denotes diminishing visual acuity.

Table 2.5. Visual acuity summary; scores are the denominator from the Snellen acuity (e.g., 20/30 →30).

		Visual Acuity					
		Distance- denominator (e.g., 20/x)			*Near Vision at 30 cm viewing distance -denominator (e.g. 20/x)*		
	N	Binocular	Best	Worst	Binocular	Best	Worst
All Participants	**27**	Range: 10-400 Mean (standard error): 57.56 (15.84) Median: 25	Range: 10-400 Mean (standard error): 56.69 (15.60) Median: 25	Range: 13-800 Mean (standard error): 96.83 (31.58) Median: 32	Range: 20-400 Mean (standard error): 57.96 (15.39) Median: 20	Range: 20-400 Mean (standard error): 55.96 (15.37) Median: 25	Range: 20-400 Mean (standard error): 87.22 (21.61) Median: 30
Desktop	***14***	Range: 13-400 Mean (standard error): 74.32 (28.31) Median: 25	Range: 16-400 Mean (standard error): 76.86 (28.32) Median: 28.5	Range: 16-800 Mean (standard error): 141.5 (57.75) Median: 51.5	Range: 20-400 Mean (standard error): 77.86 (28.18) Median: 25	Range: 20-400 Mean (standard error): 74.62 (29.35) Median: 25	Range: 20-400 Mean (standard error): 128.21(37.96) Median: 65
Controls	*2*	Range: 16 Mean (standard error): 16 (.00) Median: 16	Range: 16-20 Mean (standard error): 18 (2.00) Median: 18	Range: 16-20 Mean (standard error): 18 (2.00) Median: 18	Range: 20-25 Mean (standard error): 22.5 (2.50) Median: 22.5	Range: 20-25 Mean (standard error): 22.5 (2.50) Median: 22.5	Range: 20-25 Mean (standard error): 22.5 (2.50) Median: 22.5
AMD	*12*	Range: 13-400 Mean (standard error): 84.04 (32.32) Median: 25	Range: 16-400 Mean (standard error): 86.67 (32.32) Median: 32	Range: 16-800 Mean (standard error): 162.08 (65.79) Median: 71.5	Range: 20-400 Mean (standard error): 87.08 (32.26) Median: 32.5	Range: 25-400 Mean (standard error): 84.10 (34.10) Median: 30	Range: 25-400 Mean (standard error): 145.83 (42.30) Median: 80
Handheld	***13***	Range: 10-160 Mean (standard error): 39.50 (11.90) Median: 25	Range: 10-125 Mean (standard error): 34.96 (9.11) Median: 20	Range: 13-200 Mean (standard error): 48.73 (14.99) Median: 25	Range: 20-100 Mean (standard error): 36.54 (8.133) Median: 25	Range: 20-100 Mean (standard error): 37.31 (8.04) Median: 25	Range: 20-125 Mean (standard error): 43.08 (10.35) Median: 25
Controls	*4*	Range: 10-40 Mean (standard error): 22.75 (6.52) Median: 20.5	Range: 10-32 Mean (standard error): 20.5 (4.5) Median: 20	Range: 16-40 Mean (standard error): 25.25 (5.25) Median: 22.50	Range: 20-25 Mean (standard error): 21.25 (1.25) Median: 20	Range: 20-25 Mean (standard error): 22.5 (1.44) Median: 20	Range: 25-40 Mean (standard error): 28.75 (3.75) Median: 25
AMD	*9*	Range: 13-160 Mean (standard error): 46.94 (16.66) Median: 25	Range: 13-125 Mean (standard error): 41.39 (12.63) Median: 20	Range: 13-200 Mean (standard error): 59.17 (20.93) Median: 25	Range: 20-100 Mean (standard error): 43.33(11.15) Median: 25	Range: 20-100 Mean (standard error): 43.89 (11.05) Median: 30	Range:20-125 Mean (standard error): 49.44 (14.62) Median: 30

Color vision was assessed in terms of two different tests (the comparison of which was of interest to the NSU research team). A wide variety of clinical color vision assessment methods are available, which commonly assess the type (red-green, blue-yellow) and severity of color deficit. Red-green color deficiencies are classified as Protan or Deutan and are typically congenital. Blue-yellow color deficiencies, termed Tritan, are acquired deficiencies, and therefore more prevalent in populations with AMD.

To assess color vision, American Optical Hardy-Rand-Rittler plates (Dain 2004) were used, as well as the large panel version of the D-15 test (Uvijls, Leroy, Leys, Rouck and Kestelyn 1998; Dain 2004). A pseudoisochromatic plate test, the HRR test presents patients an object, such as a letter, number or symbol, delineated by a color difference with a background of consistent luminance shown in Figure 2.2. The outcome of the HRR is assessed as a red-green, yellow-blue color deficiency, or undefined non-congenital defect.

Figure 2.2. Example of a pseudoisochromatic plate tests, the text in the foreground is the number 16 (Source: www.colorvisiontesting.com).

Color vision was also assessed using the large panel version of the D-15 test. In this arrangement test, the participant is asked to sort 15 'test caps,' each cap a different hue, into a sequence based on their relative hue, as shown in Figure 2.3. The 15 test caps were presented to participants in a semi-random order, with the reference cap always being presented first. Similar to the HRR, the outcome from the D-15 identifies the presence of color deficiency and classifies the type of deficit. The D-15 classifies the deficiency as deutan, protan, tritan, or undefined non-congenital color deficit. Color vision was assessed for each eye independently, the scores of the various cohorts summarized in Table 2.6.

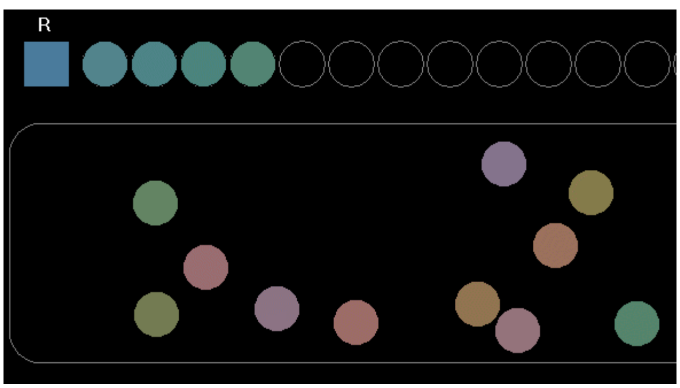


Figure 2.3. Sample color arrangement test; patients sort the colors in order of hue, similar to the D-15 test, 'R' labels the reference cap.

Contrast sensitivity, the measure of how visible an object is before it is indistinguishable from the environment, was assessed using the Pelli-Robson contrast sensitivity assessment (Pelli, Robson and Wilkins 1988). The test counts the number of letters on the Pelli-Robson Chart, shown in Figure 2.4, which can be accurately perceived from the background. Contrast sensitivity was evaluated for the left and right eyes independently in addition to binocular sensitivity. The contrast sensitivity of the participants in this study is summarized in Table 2.6.

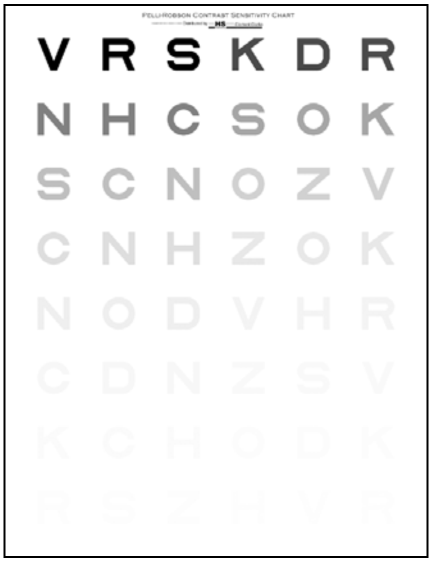


Figure 2.4. Sample Pelli-Robson contrast sensitivity test (image source: http://www.psych.nyu.edu/pelli/pellirobson).

Preferential Hyperacuity Perimeterty (PHP) was used to assess both visual function and ocular disease. This test analyzes an individual’s visual field map for distortions and blind spots that are analogous to the difference between the two stages of AMD. PHP is an automated and standardized analysis of the visual field, based on hyperacuity. Hyperacuity describes the ability to resolve the different between two relative spatial locations between stimuli. In patients with AMD, as the retinal pigment epithelium (RPE), shows elevation from the presence of drusen, these patients are more likely to experience a shift in the location of visual stimulus and/or distortion of the visible image.

Table 2.6. Summary of contrast sensitivity and color vision assessments for participants; color vision is summarized using the incidence rates for each type of deficiency type.

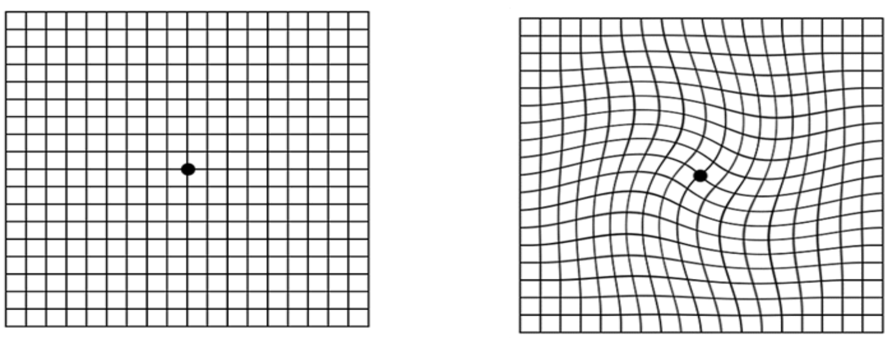
	N	Contrast Sensitivity			Color Vision	
		Binocular	*Best*	*Worst*	*D-15* *Deficiency Eye1- Eye2*	*HRR* *Deficiency Eye1- Eye2*
All Participants	**27**	Range: 11-42 Mean (standard error): 31.96 (1.53) Median: 35	Range: 11-42 Mean (standard error): 32.15 (1.53) Median: 35	Range: 6-36 Mean (standard error): 28.00 (1.74) Median: 30	None: 19 Deutan-Deutan: 1 Protan-Protan: 1 Protan-Undefined: 1 Tritan-Tritan: 1 Tritan-Undefined: 1 Tritan-None: 1 None-Undefined: 1	None: 20 BlueYellow-BlueYellow:1 RedGreen-None: 1 Undefined-Undefined: 4 None-Undefined: 1
Desktop	***14***	Range: 11-42 Mean (standard error): 29.86 (2.61) Median: 34	Range: 11-42 Mean (standard error): 29.93 (2.62) Median: 36	Range: 6-36 Mean (standard error): 25.79 (3.10) Median: 28.5	None: 8 Deutan-Deutan: 1 Protan-Protan: 1 Protan-Undefined: 1 Tritan-Tritan: 1 Tritan-Undefined: 1 None-Undefined: 1	None: 9 RedGreen-None: 1 Undefined-Undefined: 3 None-Undefined: 1
Controls	*2*	Range: 36-38 Mean (standard error): 37.0 (1.00) Median: 37	Range: 36-38 Mean (standard error): 37.0 (1.00) Median: 37	Range: 36-36 Mean (standard error): 36.0 (.00) Median: 36	None: 2	None: 2
AMD	*12*	Range: 11-42 Mean (standard error): 28.67 (2.91) Median: 28.67	Range: 11-42 Mean (standard error): 28.75 (2.93) Median: 32	Range: 6-36 Mean (standard error): 28.75 (3.38) Median: 27	None: 6 Deutan-Deutan: 1 Protan-Protan: 1 Protan-Undefined: 1 Tritan-Tritan: 1 Tritan-Undefined: 1 None-Undefined: 1	None: 7 RedGreen-None: 1 Undefined-Undefined: 3 None-Undefined: 1
Handheld	***13***	Range: 27-42 Mean (standard error): 30.38 (1.23) Median: 30	Range: 27-42 Mean (standard error): 34.54 (1.24) Median: 35	Range: 23-36 Mean (standard error): 30 (1.22) Median: 30	None: 11 Tritan-None: 1 None-Undefined: 1	None: 11 BlueYellow-BlueYellow: 1 Undefined-Undefined: 1
Controls	*4*	Range: 33-41 Mean (standard error): 37.25 (1.75) Median: 37.5	Range: 33-41 Mean (standard error): 37.25 (1.75) Median: 33	Range: 30-36 Mean (standard error): 33 (1.73) Median: 33	None: 3 Tritan-None: 1	None: 4
AMD	*9*	Range: 27-42 Mean (standard error): 32.89 (1.66) Median: 35	Range: 27-42 Mean (standard error): 33.33 (1.52) Median: 35	Range: 23-36 Mean (standard error): 29.22 (1.49) Median: 28	None: 8 None-Undefined: 1	None: 7 BlueYellow-BlueYellow: 1 Undefined-Undefined: 1

This is especially detectable when the stimulus is a straight line. In this study the visual field was assessed using the Zeiss, PreviewPHP™, which evaluates the central 14 degrees of binocular vision. The PreivewPHP™ is a computer-based method in which the participant identifies distortion of a straight line stimulus using a pen-based input display over several trials, shown in Figure 2.5. The output from the PreivewPHP™ is the identification of a visual disturbance consistent with AMD (wet form) and a graphical pattern of the visual disturbance, also illustrated in Figure 2.5 and its results summarized in Table 2.7. For our participants, the results of the PHP proved inconclusive; none of the perimetry tests denoted visual field disturbances, and one participant was unable to perform the test.

(a) (b)

Figure 2.5. Illustration of Preview PHP™ test administration (a), and visual field disturbance pattern generated (b) (image source: www.meditec.zeiss.com).

Distortions and vision loss in the central visual field associated with the macula were assessed by means of the Amsler Grid, shown in Figure 2.6(a). For this test, patients viewed the square grid with the dot in the center at about 14 inches away and using corrective frames if they typically use them for reading. The patient notes any discontinuities in the lines of the grid and/or around the perimeter of the square including holes, blurry spots, or the wavy, fuzzy, crooked lines, shown in Figure 2.6(b). For the test, left and right eyes were tested independently, as was binocular vision. The test was scored as *No Defect*, *Metapmorphisia* (e.g., distortion)*, Scotoma* (blind spot), or *E-centric viewing* (i.e. accomplished using their peripheral vision). Table 2.7 includes a synopsis of the binocular Amsler Grid Scores for the varied participant cohorts and experimental conditions.



(a) No Defect (b) Distortion

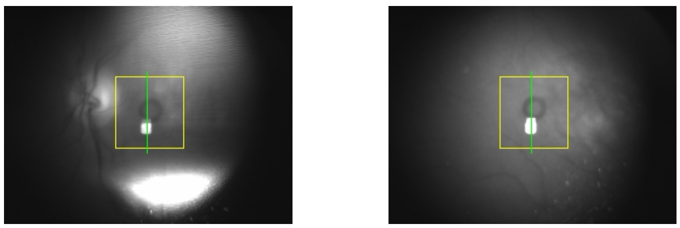
Figure 2.6. Amsler Grid Test; (a) Test grid, as presented to a patient with no defect present and (b) Test Grid as it may appear to someone with metamorphisia attributable to AMD (From http://www.macular-degeneration.org).

Central to this study was the NSU team's ophthalmic examination of the patients' eyes for the diagnosis of AMD. Accurate diagnosis of AMD is best achieved through ophthalmic examination coupled with the review of color photographs of the aspects of each eye. Numerous distinctive visible features on the eye facilitate the diagnosis and classification of the disease. The NSU optometrists visually scanned each patient's maculae for the presence of visible drusen – discrete yellowish-white spots on the retina. In addition they examined the state of the retinal pigment epithelium (RPE), a single layer of cells between the retina and the underlying blood vessels, for leakage. This was accomplished through the direct examination of patients' eyes, and in some instances further review of photographs of the retina, or fundas images, such those provided in Figure 2.7.

Several classification systems have been used to grade the severity of AMD. The present study employed a method introduced in 1989 (Bressler, Bressler, West, Fine and Taylor 1989), which grades severity level of the disease on a scale from 0 (no disease) to 4 (most severe) based on the amount of drusen, their distribution on the macula and observed condition of the RPE. Grade 4, the most severe or final stage is assigned to those cases in which there is geographic atrophy of the RPE or new blood vessel grown below the retina, associated with the wet form of AMD.

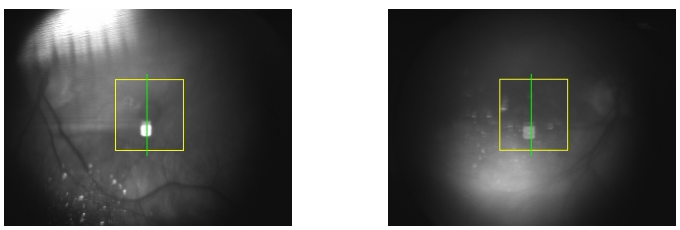
Left Eye Right Eye

(a) P21



Left Eye Right Eye

(b)P17



Left Eye ` Right Eye

(c) P5

Figure 2.7. Sample fundas images for three participants.

The Chesapeake Bay Study scoring system was intended for patients 30 years or older, and allots the following scores to the condition of the eye:

- *Grade 1*: at least 5 small drusen within 1500μm of the foveal center, or at least 10 small drusen between 1500μm and 3000μm of the foveal center
- *Grade 2*: Many small drusen ~20 or more, within 1500μm of the foveal center
- *Grade 3*: Eyes with large confluent drusen or eyes with focal hyperpigmentation of the RPE
- *Grade 4*: Geographic atrophy of the RPE or exudative changes

  (Bird et al., 1995)

In the present study, Grade 0 was given to those eyes without any drusen, or fewer than five drusen. Each eye was graded independently using the scoring system. Participants with Grade 0 in both eyes were identified as part of the Control group. In addition the NSU team rated the type of AMD present in each eye as 'Wet' or 'Dry' according to any visible leakage discerned on the RPE. An overview of the prevalence of the type of AMD and the severity levels observed in the participants' eyes are presented in Table 2.7.

As with the participant characteristic demographics, statistical analyses were performed to identify any potentially significant differences between those participants in the desktop and handheld conditions, as well as between the AMD and control cohorts. These statistical analyses are summarized Table 2.8. The Mann-Whitney non-parametric test confirmed there were no statistically significant differences between the handheld and desktop cohorts for binocular distance visual acuity ($p = 550$), near visual acuity ($p = .280$) or contrast sensitivity ($p = .458$.), or worst eye AMD severity level ($p = .239$). Mann-Whitney comparisons also showed no significant differences in near visual acuity, distance visual acuity, or contrast sensitivity between the controls and AMD cohorts within either the desktop or the handheld PC experimental conditions ($p > .132$).

Chi Squared Tests examined differences in AMD Score, HRR and D-15 color tests results, and the outcome of the Amsler grid test (binocular). The test detected no significant differences between the handheld and desktop cohorts, nor between Controls and AMD groups (within each condition on HRR, D-15, or Amsler ($p > .19$)).

There were no statistical differences between the desktop and handheld conditions for AMD Score ($p = .304$).

Not surprisingly, significant differences were detected both within desktop and handheld conditions between the controls and AMD for AMD Score (X^2 (3, N= 27)= 8.73 p = .003; X^2 (3, N= 27)= 14.0 p = .003) and AMD Severity Level (Mann-Whitney U (N = 14) <.001, Z = -.231, p =.022; Mann-Whitney U (N = 13)< .001, Z = -2.90, p =.003). As these two measures, AMD severity level, and AMD Score are indicative of the presence of the disease, the results of these statistical tests confirm the absence of AMD in the control group.

Table 2.7. Summary of assessments of AMD diagnoses; Amsler Grid, Type of AMD and Severity Score columns summarize the incidence rate of the type or level.

	N	**Visual Field**		**Type of AMD** *(Eye 1-Eye 2)*	**AMD Severity Score** *(Eye 1-Eye 2)*
		PHP	Amsler Grid *(binocular)*		
All Participants	**27**	No disruption: 27 Unable to perform: 1	No disturbance: 21 Eccentric viewing: 2 Metamorphisia: 2 Scatoma: 2	No AMD: 6 Dry-Dry:17 None-Dry:2 Wet-Wet: 2	0-0: 6 1-0: 2 1-1: 9 1-2: 1 2-2: 1 3-3: 3 4-4: 5
Desktop	***14***	No disturbance: 13 Unable to perform: 1	No disturbance: 9 Eccentric viewing (both eyes): 1 Metamorphisia (both eyes): 2 Scatoma (both eyes): 1	No AMD: 2 Dry-Dry: 8 None-Dry:2 Wet-Wet: 2	0-0: 2 1-0: 2 1-1: 4 3-3: 2 4-4: 4
Controls	*2*	No disturbance: 2	No disturbance: 2	No AMD: 2	0-0: 2
AMD	*12*	No disturbance: 11 Unable to perform: 1	No disturbance: 7 Eccentric viewing (both eyes): 1 Metamorphisia (both eyes): 2 Scatoma (both eyes): 2	Dry-Dry: 8 None-Dry:2 Wet-Wet: 2	1-0: 2 1-1: 4 3-3: 2 4-4: 4
Handheld	***13***	No disturbance: 13	No disturbance: 12 Eccentric viewing: 1	No AMD: 4 Dry-Dry: 9	0-0: 4 1-1: 5 1-2: 1 2-2: 1 3-3: 1 4-4: 1
Controls	*4*	No disturbance: 4	No disturbance: 4	No AMD: 4	0-0: 4
AMD	*9*	No disturbance: 9	No disturbance: 8 Eccentric viewing (both eyes): 1	No AMD: 6 AMD type (Eye1-Eye2): Dry-Dry: 9	1-1: 5 1-2: 1 2-2: 1 3-3: 1 4-4: 1

Table 2.8. Summary of statistical analyses of visual factors.

Visual Assessment	Statistical Test	Comparison		
		Desktop vs. Handheld	*Desktop: Control vs. AMD*	*Handheld: Control vs. AMD*
Distance Visual Acuity (binocular)	Mann-Whitney	Mann-Whitney U (N = 27)= 78.0, Z = -.639, p =.550	Mann-Whitney U (N = 14) = 3.00, Z = -1.67, p =.132	Mann-Whitney U (N = 13) = 12.5, Z = -.859, p =.414
Near Visual Acuity (binocular)		Mann-Whitney U (N = 27)= 68.5, Z = -1.13, p =.280	Mann-Whitney U (N = 14) = 6.00, Z = -1.12, p =.352	Mann-Whitney U (N = 13)= 8.50, Z = -1.55, p =.148
Constrast Sensitivity (binocular)		Mann-Whitney U (N = 27)= 75.0, Z = -.782, p =.458	Mann-Whitney U (N = 14) = 3.00, Z = -1.65, p =.132	Mann-Whitney U (N = 13)= 8.5, Z = -1.48, p =.148
AMD SEVERITY SCORE (worst eye severity)		Mann-Whitney U (N = 27)= 66.0 Z = -1.27, p =.239	Mann-Whitney U (N = 14) <.001, Z = -.231, p =.022*	Mann-Whitney U (N = 13)< .001, Z = -2.90, p =.003*
AMD TYPE	Chi-Squared	X^2 (3, N= 27) = 3.64, p = .304	X^2 (3, N= 27) = 14.0, p = .003*	X^2 (3, N= 27) = 8.73 p = .003*
HRR COLOR DEFICIENCY*		X^2 (2, N= 27) = .585 p = .444	X^2 (3, N= 27) = .117, p = .733	X^2 (3, N= 27) = .037, p = .848
D-15 COLOR DEFICIENCY*		X^2 (2, N= 27) = 1.30, p = .254	X^2 (3, N= 27) =.304, p = .581	X^2 (3, N= 27) <.001, p = 1.00
Amsler Grid Test		X^2 (3, N= 27) = 1.66, p = .198	X^2 (3, N= 27) <.001, p = 1.00	X^2 (3, N= 27) <.001, p = 1.00

*In order to meet the assumptions for Chi-Squared distribution, color deficiency for both HRR and D-15 were recoded into three levels: 0 = no deficiency; 1 = deficiency in at least one eye.

Experimental Methods & Procedures

Upon arrival to the clinic, participants were briefed on the purpose of the study and informed consent was obtained following the protocols for both NSU and the Georgia Institute of Technology (Georgia Tech). Next, all ophthalmic tests were performed at the NOVA clinic, excluding those tests involving dilation of the eyes (as to not interfere with computer performance and the measure of pupillary response to task performance). When necessary, participants were provided with temporary frames outfitted with the appropriate corrective lenses to enable use of their best-corrected vision for the experimental tasks. Participants then worked with the team from Georgia Tech on a series of computer-based tasks and surveys. Finally, the remainder of ophthalmic examination was carried out. On average, participants were at the clinic for three hours.

In addition to the clinically acquired measures of visual function and visual health, two additional vision-related tests were administered, classified in this study as non-clinically acquired assessments of visual function. The first is the Visual Functioning Questionnaire (VFQ-25), a subjective assessment of visual functioning created by the National Eye Institute (Mangione, Berry, Spritzer et al. 1988; Mangione, Lee, Gutierrez et al. 2001). This interview administered questionnaire produces several sub scores, each based on questions concerning the level of difficulty incurred in activities of daily living, related to visual function. In previous studies when clinical vision diagnostic tests (e.g. tests of visual acuity, contrast sensitivity, and color perception) failed to reveal significant differences in sensory capabilities (see Table 2.8), the results from the NEI VFQ-25 shed light on some of the underlying group differences (Jacko, Barnard, Kongnakorn et al. 2004; Jacko, Moloney, Kongnakorn et al. 2005), and the VFQ-25 has been shown indicative of performance differences in previous experiments (Leonard, Edwards and Jacko 2005). The subscales used in the VFQ are comprise questions concerning:

- General health
- General vision
- Ocular pain
- Near activities
- Distance activities
- Vision Related
    - Social functioning
    - Mental health
    - Vision Related
    - Role difficulties
    - Dependency
- Driving
- Color vision
- Peripheral vision

The second additional non-clinical assessment of visual function is a visual attention test. Visual attention is a function of both cognitive and physical sensory faculties, and is not captured via measures of ocular health. Visual attention is interesting in characterizing HCI for the population of users with visual impairments. An individual with visual impairments may or may not experience detriments to their visual attention capacity, as they can acquire skills to cope with their visual dysfunction (e.g., learning new scanning strategies to acquire visual information). Visual attention is typically assessed for individuals with neurological or brain damage (e.g. stroke) who may experience hemispheric visual loss, or cognitive deficits (e.g., dyslexia). The measurement of visual attention has not been examined in previous HCI research for individuals with ocular diagnoses.

The Visual Search and Attention Test (VSAT) (Trenerry, Crosson, DeBoe and Leber 1990) is a visual cancellation test to assess an individual's capacity for visual search and attention. These are tests of sustained attention, with four repetitions. The testing materials contain normative data for a variety of age groups for statistical comparisons. Two additional assessments of the participants, summarized by Table 2.9, were collected before the completion of the computer task. This included evaluations of mental health, physical health and manual dexterity, all of which have been shown, in previous studies, to impact computer interaction for older adults who have visual impairments (Emery, Jacko, Kongnakorn et al. 2001; Jacko, Scott, Sainfort et al. 2002; Emery, Edwards et al. 2003; Jacko, Moloney et al. 2005). A photograph of the Purdue Pegboard is given in Figure 2.8.

Figure 2.8. Purdue Pegboard (image source: wisdomking.com).

Table 2.9. Cohort variables.

Characteristic	Test (metric)
Manual dexterity	A count of the average number of pins inserted into small holes in a board over three, 30 second trials, from 0 (worst) to 30 (best) (Tiffin and Asher 1948)
Mental & Physical Health (self reported)	Short Form 12 Health Survey (SF-12™) Health Survey, which yields general composite scores of mental and physical health (Ware, Kosinski and Keller 1995).

2.3. Experimental Task

This dissertation investigated, for a complex drag and drop task (targets in the presence of distracters), the measures of visual function and visual health on the different components of the interaction (e.g., visual search, target selection, target movement and drop) on both the handheld and desktop computer tasks. The following factors were examined in detail in association with characteristics of visual dysfunction:

- User interactions in the presence of visual and auditory interface enhancements
    - Screen density (icon size, set size, spacing)
    - Presence and absence of an auditory icon for feedback
- User interactions in different contexts
    - Device, handheld PC and desktop PC

The icons employed in this experiment followed the *simple* and *good* guidelines put forth by (Byrne 1993). While participants were screened for computer experience, it is presumptuous to anticipate older adults have a high level of familiarity with the typical office application icons, such as Microsoft Word® icons and Windows Explorer® file folders used in previous experiments (Jacko, Barreto, Chu et al. 2000; Jacko, Scott, Sainfort et al. 2003; Vitense, Jacko and Emery 2003). So, while Sears and Jacko, in their framework for icons, list the office suite icons as the most common (Sears, Jacko, Brewer and Robelo 1998), it could be argued that this is dependent on the users' computer experience and exposure to these applications.

The majority of older adults' computer experience is derived from 1) Internet use, 2) email, and 3) games (Emery, Edwards et al. 2003; Jacko, Scott et al. 2003; Jacko, Moloney et al. 2005). In fact, in our previous field studies, some participants first claimed no knowledge of computers, but upon further dialogue spoke excitedly about playing games such as Solitaire or Minesweeper™ almost on a daily basis on their machines.

To limit the effects of disparities in participants' comfort levels and familiarity with the icons used in this experiment, images of playing cards, similar to those found in the solitaire game included in the Windows® platform (see Figure 2.9), served as the icons. The motivation for the use of card icons was to equalize the familiarity and comfort level with the icons between participants and to generate experimental results generalizable to future populations of older adults who are likely to be more homogeneously equivalent in terms of computer experience and comfort levels. The familiarity of the participants with the card symbols is much more likely. Moreover, an increasingly large number of older adults play card games on a regular basis as it has been shown to mitigate effects of aging and dementia (Coyle 2003).

The playing card icons are analogous to the icons recommended by Byrne (1993); simple, graphic icons best discriminated by as few features as possible. Furthermore, in later work by Everett and Byrne (2004) the authors used good, fair, and poor icons. Simple shapes and colors identified the "good" icons. Each playing card icon contained both a graphic and textual components. The card icons used were easily detected and highly familiar to the participants in the study in order to control for the effects of learning, icon design. This facilitated a higher degree of experimental control in the isolation of the impact of varied levels of spacing, size, set size, feedback, and interaction device.

Figure 2.10. illustrates the commonalities between Byrne's *simple* icons (not to be confused with the simple versus complex task environment), Everett and Byrne's *good* icons, and the playing card icons proposed for use in this dissertation. It should also be noted that these studies incorporated text labels beneath the icon

graphic. According to Byrne (1993) “Effective icons must be simple and easily discriminated. Simple icons are more effective with larger set sizes, allow effective use of icon knowledge, are less affected by a lack of file name knowledge, and are especially effective when they are unique to the display; with simple icons, there is reason to accept the design assumption that icon pictures make finding files easier,” (p.452).

The playing card icons used in the study were numbered 2 through 9, to enable consistency in visual search (no aces, queens, kings or jacks, to exclude cards with letters instead of numbers, and those with detailed face card illustrations). All four suits were represented, hearts, diamonds, clubs and spades (e.g. ♥ ♦ ♣ ♠), in traditional red and black color-coding and the background of each card was white.

As shown in Figure 2.11, the card icons had two components, the graphic representing the suit, and the text identifying the number. In lieu of folder icons used as the target destination of icon movement in previous work (e.g., Vitense, Jacko and Emery 2002; Vitense, Jacko et al. 2003; e.g., Jacko, Moloney et al. 2005) icons representing stacks of upside down playing cards, called *drop piles*, were used to carry out the card metaphor. On each trial, four drop piles were always present in the left-most column of the screen, one card for each suit, labeled with the appropriate suit symbol as illustrated in shown in Figure 2.11.

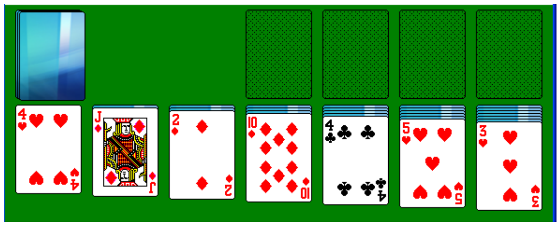
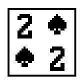
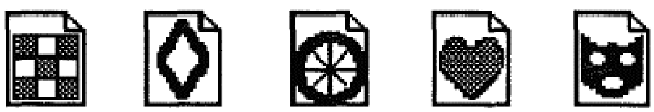


Figure 2.9. Windows® Solitaire card game.



Playing card icon



Simple icons, used by Byrne (1998)

Good, fair and poor icons used
by Everett and Byrne (2004)

Figure 2.10. Juxtaposition of card icons with *Simple* and *Good* icons (Byrne, 1993; Everett and Byrne 2004).

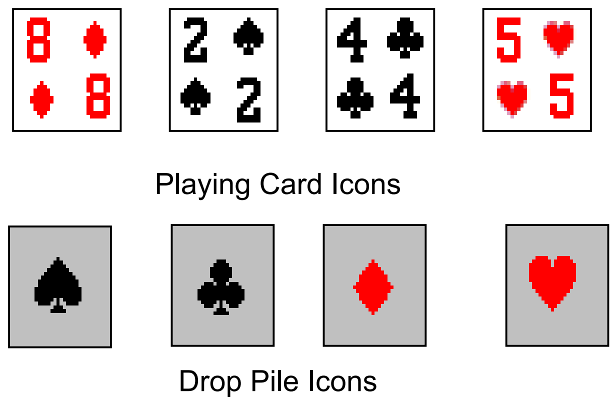


Figure 2.11. Sample icons.

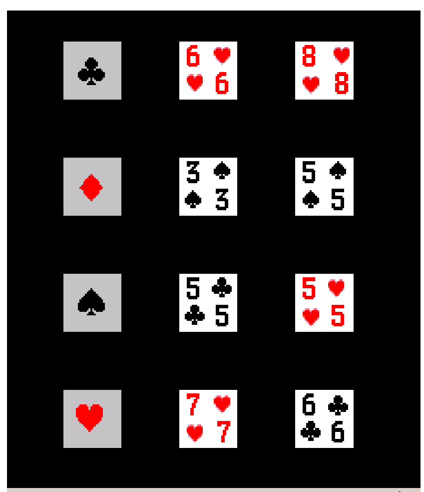


Figure 2.12. Trial screen with a grid of card icons and a column of drop piles on the left.

At the start of each trial the user was presented with a grid of icons, its size dependent on the number of icons, spacing and icon sizes employed. Figure 2.12 illustrates the layout of the display for both the handheld and desktop. In each experimental condition, the leftmost column of the grid was comprised of the drop piles, the target destination for the card icons to be relocated to according to the matching suit. This location was chosen as it is consistent with the standard location of folders in the default Windows® desktop setting *Arrange icons by . . . file type* and consistent with the location of folders in Windows® Explorer file management program. The sizes of the icons, set size, and spacing were derived from the work of both Jacko and colleagues (1999) and Byrne and Everett (Byrne 1993; Everett and Byrne 2004). The contrast of the display was a high contrast, using a black background with white card icons, consistent with the findings of Jacko and colleagues (Jacko, Rosa et al. 1999; Scott, Feuer and Jacko 2002).

Under conditions that provided auditory feedback, an auditory icon was used that provided the user with a ‘sucking’ or scraping sound whenever a card (selected with the mouse or stylus) entered in position for an acceptable drop. That is, the sucking sound signaled to the participant that if their finger released the left mouse button, or they lifted the stylus away from the screen at the moment the feedback at the time of the sound, the icon would ‘drop into’ the pile. This sound employed was consistent with that used in previous experiments of multimodal feedback and drag and drop tasks for users with visual impairments (e.g., Jacko, Scott et al. 2002). A purplish highlight consistent with the standard Windows® desktop, provided feedback of accurate positioning of the card, as used in previous studies (Vitense, Jacko et al. 2002; Emery, Edwards et al. 2003; Jacko, Moloney et al. 2005). Note that neither the sound nor the highlight informed the user if they were positioned over the *correct* drop pile, only that they were in position for a card drop into any one of the four piles.

For both the handheld and the desktop conditions, the trials were grouped into two sets, AF_1 and AF_0. All trials with auditory feedback present (AF_1) were performed either entirely first or entirely second, this order randomly assigned to each participant. The grouping of the AF conditions was intended to enable the participants to make appropriate use of the feedback and mitigate any carry over effects if one trial had auditory and then the next did not. Participants were made aware at what point in time they should begin to anticipate the auditory feedback, as well as when they should cease to anticipate. Within each set of trials, AF_1 and AF_0, the order of exposure to the remaining experimental conditions was fully randomized between participants.

Participants were briefed on the purpose of the experiment and the task was demonstrated for them, including the different experimental levels. In the handheld experiment, participants were all novices to the handheld PC. Experimenters therefore, provided a brief orientation on the use of the stylus. Next, for both conditions, the volume was adjusted so that auditory feedback was easily perceptible by each participant. Participants then performed a series of self-terminating practice trials.

On each trial participants were verbally instructed to: 1) locate a specific target card amongst a grid of several distracter card icons of different numbers/suits; 2) select the target using the stylus or mouse; and 3) drag it to the card pile on the left-hand side of the display which matched its suit and drop the card into this pile.

The instructions of target card were given to the participants with a blank screen present on the display, which only had a button labelled ‘proceed’ consistently placed at the bottom-center of the display. Before they were allowed to select the proceed button, the participants were required to verbally repeat back the target card icon name. This provided confirmation that they were, in fact, looking for the correct target card. The participants practiced this exchange with the experimenter and the practice trials several times before the experimental trials commenced. Participants were directed to work as quickly and accurately as possible through each trial, and that the target icon for each trial would always be somewhere on the display.

For both the desktop and handheld conditions, the conditions within AF present and AF absent were completely randomized. The order of exposure to these two sets was random across participants. The location of the target card and the drop piles and the specific collection of distracter card icons and their placement were all randomly assigned for each trial across participants. The target card for each trial was consistent between participants for simplification of experimental protocol. While participants searched for the same target cards at trial 1, 2, and so on, the conditions under which they sought that icon were randomly ordered to mitigate any specific impact of the card number or suit.

All participants were fitted with the ASL-501 head-mounted eye tracker with eye-head integration, outfitted to track their eye with the best-corrected vision and to record both pupillary response and eye movements. The eye tracking system was then calibrated to the participant using a standard protocol (recommended by the manufacturer). In addition to the recording of eye movements, participants' reactions to the task were video recorded for future cording and analyses.

2.4. Experimental Design

Software for the experimental task was developed for both the handheld PC and the desktop computer task environments. That said, the interface type (computer versus handheld) served as a between subject variable. Because of limited screen real estate availability on the handheld PC, the full set of conditions on the desktop could not be fully replicated for the handheld (e.g., the number of levels of set size, spacing, and icon size were limited on the handheld). However, the experimental levels present on the handheld were wholly replicated on the desktop interface to isolate the effects handheld computer interaction apart from the identified interface permutations. This section details the experimental design and equipment setup for the desktop and handheld PC conditions separately.

Desktop PC Condition

The desktop software was developed in Visual Basic, and ran on an IBM compatible computer with an Intel Pentium III 935 MHz processor and 512 Mbytes RAM. A 20” viewable flat panel display was used, and the resolution set to 1024x768 with 15-Bit Color. Participants were seated approximately 40 cm viewing distance from the display and used an optical mouse as the input device. Figure 2.13 illustrates the experimental environment for the desktop condition.

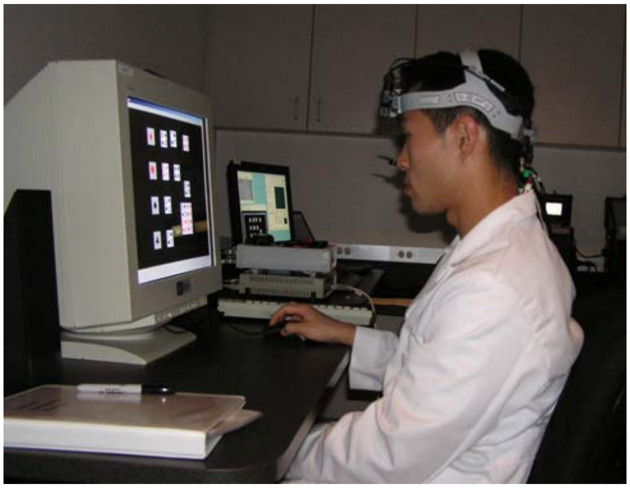


Figure 2.13. A member of the experimental team demonstrates the desktop computer experimental environment, wearing the head mounted eye tracker

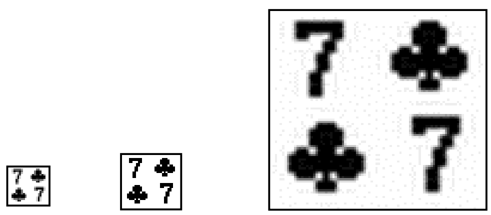
The desktop condition employed a 3x3x3x2 repeated measures design. In total there were 54 experimental permutations for the desktop, and users performed 3 repetitions of each permutation for a total of 162 trials. The independent variables controlled in this experiment included:

Icon Size (ISz)

ISz was based on icon sizes employed in real world applications and empirical work by Jacko and colleagues. The sizing was applied to both the card icons and the drop piles. ISz was investigated at three levels, the relative sizes shown in Figure 2.14.

- ISz_1: The physical size of standard icons on the handheld PC, 7mm each side, 22x22 pixels on this display.

- ISz_2: The standard windows icons size of 32x32 pixels (active area) (Microsoft Developers Network 2004), on this display and given the resolution these icons appeared 12mm each side.
- ISz_3: The icon size recommended for individuals with visual impairments, 36.8 mm (diagonal distance in the active area) 90x90 pixels on this display, per the work of Jacko and colleagues (Jacko, Rosa et al. 1999; Jacko, Moloney et al. 2005).



ISz_1 ISz_2 ISz_3

Figure 2.14. Relative sizes of icons for the desktop condition, not shown at actual size, but relative difference is accurate.

Icon spacing (ISp)

As investigated by Everett and Byrne (2004), ISp is the distance between the mid point of the borders of two icons, and consisted of three levels in the desktop condition. Spacing between icons was applied between all icons, card and drop piles. Spacing is commonly measured relative to the size of the icon, as follows:

- ISp_1: ¼ icon width;
- ISp_2: ½ icon width; and

- ISp_3: 1 icon width.

Set size (SS)

SS was based on the investigations of Everett and Byrne (2004) and Jacko et al. (1999), and the even distribution of the icons amongst the four rows created by the drop pile icons. Each of the three SS levels also contained the column of four drop pile icons. The levels included:

- o SS_1: 4 card icons (1 column card icons, 1 column drop piles);
- o SS_2: 8 card icons 4 drop piles (2 columns card icons, 1 column drop piles); and
- o SS_3: 12 card icons, 4 drop piles (3 columns card icons, 1 column drop piles).

It should be noted that Everett and Byrne investigated set sizes of 6, 12, 18, and 24 and observed differences based on set size; Jacko investigated 2, 3, 4, 5, and 6, but did not observe any main effects due to set size, asserting it was due to the relatively small set sizes used.)

Auditory feedback (AF)

The sound used was an auditory icon that provides the user with a ‘sucking’ or scraping sound when in position for an acceptable drop. This sound is consistent with that used in previous experiments of multimodal feedback and drag and drop tasks for users with visual impairments (e.g., Jacko, Scott et al. 2002). Auditory feedback had two levels:

- o AF_1: Present; and
- o AF_0: Absent

Handheld PC Condition

Participants interacted with a Dell Axim X30 Pocket PC. The handheld display was touch-sensitive LCD, measuring 3.5 inches diagonal. Resolution was set to 240x320 at 16-bit color. The device was secured to an inclined platform during the task to accommodate the collection of eye movement data shown in Figure 2.15. Participants were seated a comfortable viewing distance from the handheld, and allowed to adjust the seating for their own comfort during the task.

Figure 2.15. Experimental configuration including screen shot of the task (not actual size).

The experimental conditions for the handheld PC were constrained by its small display size. As such, icon size (IS) was held constant throughout this condition at the standard handheld PC Icon size, 32x32 (7mmx7mm). The AF was implemented at levels consistent with the desktop condition, as were SS and ISp. The handheld condition for the experiment was therefore a 3x3x2 repeated measures design, for a total of 18 permutations. Participants performed the task on the handheld 9 times per experimental permutation for a total of 162 trials.

Dependent Variables

Several dependent variables were collected to comprehensively characterize the interactions of the participants. Table 2.10 provides a classification of these measures. A snapshot of the experimental layout is provided in Table 2.11 for the desktop PC condition and Table 2.12 for the handheld PC condition.

Table 2.10. Dependent task-related measures.

Performance Measures (per condition, per subject)		**Physiological Measures** (across all trials, per subject)		**Subjective Measures** (across all trials, per subject)
		Eye Tracking		
Measures of Efficiency	Accuracy Measures	*Visual Processing*	*Mental Workload*	
Trial time (ms)	Missed opportunities: Icon acquisition* (count)	Fixation duration	Pupillary change	Post task survey: Interface preferences,
Visual search time (ms)	Task axis crossings (count)	Saccade duration		
Movement time (dragging time) (ms)	Movement direction change (count)	Fixation to saccade duration ratio		
Final target highlight time (ms)	Accidental drops (count)			
Total target highlight time (ms)	Missed opportunities: Over no drop (count)			
Dragging distance (pixels)**				
Movement variability (pixels)				
Movement error: pixels				

* Indicates measure was only taken in the desktop condition
**Indicates measure was only taken in handheld condition

Table 2.11. Sample experimental layout for the desktop computer condition permutations.

	AF_1																										
	ISz_1									ISz_2									ISz_3								
	ISp_1			ISp_2			ISp_3			ISp_1			ISp_2			ISp_3			ISp_1			ISp_2			ISp_3		
	SS_1	SS_2	SS_3	SS_1	SS_2	SS_3	SS_1	SS_2	SS_3	SS_1	SS_2	SS_3	SS_1	SS_2	SS_3	SS_1	SS_2	SS_3	SS_1	SS_2	SS_3	SS_1	SS_2	SS_3	SS_1	SS_2	SS_3
Condition	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27

	AF_2																										
	ISz_1									ISz_2									ISz_3								
	ISp_1			ISp_2			ISp_3			ISp_1			ISp_2			ISp_3			ISp_1			ISp_2			ISp_3		
	SS_1	SS_2	SS_3	SS_1	SS_2	SS_3	SS_1	SS_2	SS_3	SS_1	SS_2	SS_3	SS_1	SS_2	SS_3	SS_1	SS_2	SS_3	SS_1	SS_2	SS_3	SS_1	SS_2	SS_3	SS_1	SS_2	SS_3
Condition	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54

Table 2.12. Sample experimental layout e for the handheld computer condition permutations.

	AF_1								
	ISp_1			ISp_2			ISp_3		
	SS_1	SS_2	SS_3	SS_1	SS_2	SS_3	SS_1	SS_2	SS_3
Condition	1	2	3	4	5	6	7	8	9

	AF_2								
	ISp_1			ISp_2			ISp_3		
	SS_1	SS_2	SS_3	SS_1	SS_2	SS_3	SS_1	SS_2	SS_3
Condition	10	11	12	13	14	15	16	17	18

Performance Measures

The performance measures of efficiency and accuracy (e.g., errors) were collected in the background by the software designed to provide the user the visual search, and card icon environment. Performance measures constituted assessments of efficiency and accuracy. Tables 2.13 and 2.14 provide detailed summaries of the measures. It should be noted that measures of the accuracy icon selection were not collected on the handheld. This is because the stylus does not make contact with the screen until an icon is selected and dragged to the target destination. Figures 2.16-2.18 provide additional details on some of the more complex metrics.

Table 2.13. Definitions of measures of efficiency.

Efficiency Measure	Definition
Trial Time (ms)	Time to complete visual search and drag and drop of a single card icon, measured from the time of display onset, until the release of a card icon into any pile (not necessarily accurate).
Visual Search Time (ms)	The time from the appearance of the screen of icons until the cursor was placed in the final selected card icon's active area (that is the card that is ultimately released into a drop pile).
Movement Time (dragging time) (ms)	The amount of time from when the mouse cursor or stylus selected a card icon (via a click of the left mouse button when the cursor was within an active card icon region or the stylus' contact with the portion of the display within an active card area) until it was successfully dropped into a pile. This measurement was taken only for the card ultimately released into a pile.
Final Target Highlight Time (ms)	The amount of time the icon was in a correct position to be dropped before it was successfully released into a drop pile. Also indicative of the amount of time that feedback (auditory and visual, or just visual) was provided to the user (Emery et al., 2003; Jacko, Moloney et al., 2005), on the final positioning of the icon, pre-release.
Total Target Highlight Time (ms)	Total amount of time feedback was provided to the users before dropping the card into a pile. This accounts for every approach to the pile, not just the final/ accurate approach (Emery et al., 2003; Jacko, Moloney et al., 2005), and reports the total amount of time feedback was provided to the users over the course of a single trial.
Dragging Distance (pixels)**	The number of pixels traveled by the mouse/stylus while holding onto the dropped card icon over the course of each trial.
Movement Variability (pixels)	The standard deviation in the distances of the path taken from the task axis from the mean for approach to the drop pile during the movement of the card ultimately dropped (MacKenzie et al., 2001).
Movement Error (pixels)	The average deviation of the sample points from the task axis (absolute distance) for the approach to the drop pile with the card ultimately dropped (MacKenzie et al., 2001).

**Measure only taken on handheld PC

Table 2.14. Definitions of measures of accuracy.

Accuracy Measure	Definition
Missed opportunities: Icon acquisition* (count)	The number of times the cursor entered the ‘selectable’ area of the correct icon (Jacko et al., 2005).
Missed opportunities: Over no drop (count)	The number of times the cursor, with the card, entered the acceptable area for drop over the pile (Jacko et al., 2005).
Accidental drops	The number of times the card was dropped prematurely on its approach to the drop pile.
Task axis crossing (count)	The number of crossings on the approach to the card pile during movement of the card to the pile.
Movement direction change (count)	The number of times the pointer’s direction changes, relative to the task axis, on dragging the card to the pile (MacKenzie et al., 2001).

*Indicates measures not collected for the handheld PC interaction

** All measures ere recorded relative to the card icon ultimately released into a pile, regardless of accuracy of card icon or pile, accuracy of drop and pile were also tracked, reported in the summaries of chapters 3 and 4.

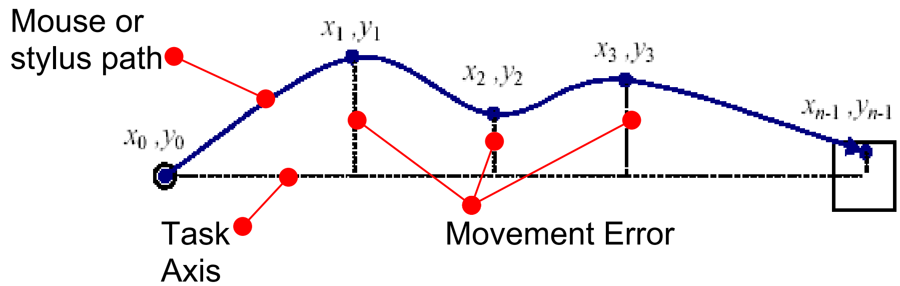


Figure 2.16. Illustration of task axis versus movement path, with measurement of movement error.

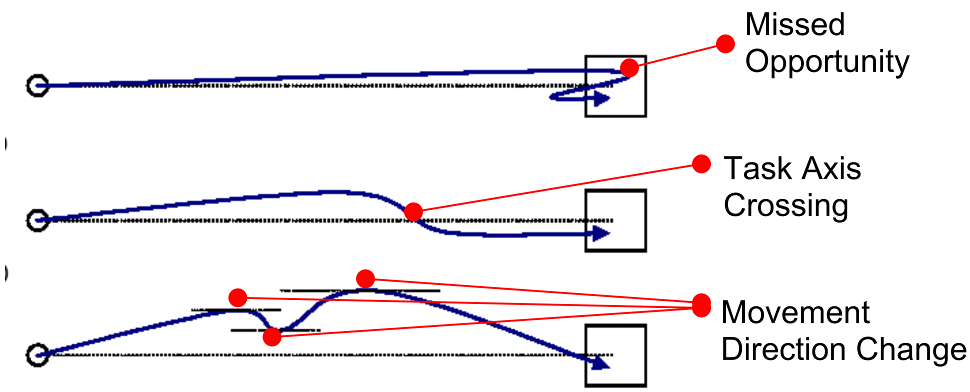


Figure 2.17. Illustration of missed opportunities, task axis crossings and movement direction change measures of efficiency.

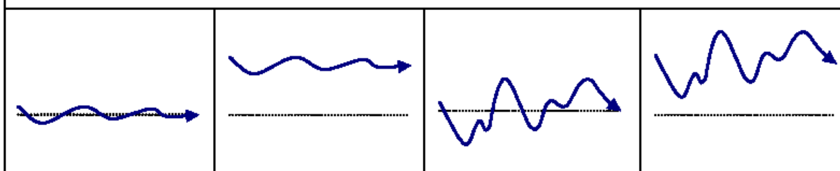
	Movement Responses			
				
Movement Variability	Low	Low	High	High

Figure 2.18. Illustration of movement variability metric.

Physiological Measures of Task Interactions

Physiological measures can provide an insightful view of the impact of interaction on a user's covert processes, particularly information processing (see Andreassi 2000). Table 2.15 provides an overview of the physiological measures collected during this experiment. Because physiological measures are highly complex and highly susceptible to noise from both the subject and test environment special care must be taken in what inferences are made from these measures. That said, the following section provides justification and explanation for the inclusion of physiological measure of workload and visual processing in the context of this research.

Table 2.15. Definitions of physiological measures.

Physiological Measure	Definition
Mental workload: Pupillary response (Continuous function, measured at 60 hertz)	The deviation in pupil diameter, per experimental condition can be indicative of the amount of workload experienced (Boff and Lincoln 1988).
Information processing: Saccades and fixations	The number of fixations and saccades can indicate the magnitude of workload imposed by information processing (May, Kennedy, Williams, Dunlap and Brannan 1990).
Eye movements: Saccade to fixation distribution	Visual search and efficiency of visual search can be measured by the ratio of saccades to fixation, and their durations and lengths relative to the display conditions (Backs and Walrath 1992).

Mental Workload

Mental workload has been defined as a measure of the demands that information processing imposes upon a human in a system (Sanders and McCormick 1993). While a clear-cut definition of mental workload lacks in the literature, the concept of mental workload has been explored to greater lengths in the research community (Sheridan 2002). Several authors have helped to vindicate the concept of workload. In their often-cited work, Gopher & Donchin (1986) portrayed workload as the cost that users incur due to information processing. The human operator has a limited capacity to process and respond to information (Kahneman 1973; Tsang 1997). The definition of workload in terms of information processing also brings to surface issues of attention. In both workload and attention, it is suggested that a limit on either reflects the organization of a person’s central processing system.

Information theory stresses that a communication channel is defined by its capacity to send information between the sender and receiver, and quantified by attributes that make up the ‘channel capacity’ (Gopher and Donchin 1986). A more articulated definition of mental workload emerges from this theoretical perspective. Workload is an indication of the difference between the information capacity of the operator and that capacity required for criterion task performance (Gopher and Donchin 1986; O'Donnell and Eggemeier 1986). Mental workload is particularly intriguing for a population of users with visual impairments, because of the apparent limitations the bandwidth of their visual sensory channel.

Wickens (1980) introduced the concept that people have separate processing resources, each with a limited capacity. Tasks that demand the same resources will result in performance decay if they cannot effectively allocate information processing resources effectively. This is known as multiple resource theory. The three dimensions of information processing include 1) processing stages, 2) processing modalities, and 3) processing codes (Wickens and Lui 1988). These theories of information processing lend a multidimensional quality to mental workload, making it a ‘vector quantity’ associated with a number of objectives (Derrick 1988).

Workload assessment serves a diagnostic or task characteristic function (Wickens and Yeh 1982). As a diagnostic appraisal, a workload measure falling above (or below) a predetermined threshold may forecast performance decrements or operator errors (e.g., Wierwille, Tijerina, Kiger et al. 1996). Figure 2.19 presents an adaptation of a popular figure in fundamental workload literature that demonstrates the predictive power of workload (O'Donnell and Eggemeier 1986; Sheridan 2002). At the onset of capacity overload, the there is an immediate decline in performance. Not represented in this figure is the case of extremely low operator workload, which also tends to exhibit drops in performance. The application of workload measures for task characterization can influence mission critical decisions in task allocation, as well as inform decisions in selecting competing designs, or operators for a particular task (Wilson 2001; Brewster 2002; Vitense, Jacko et al. 2003).

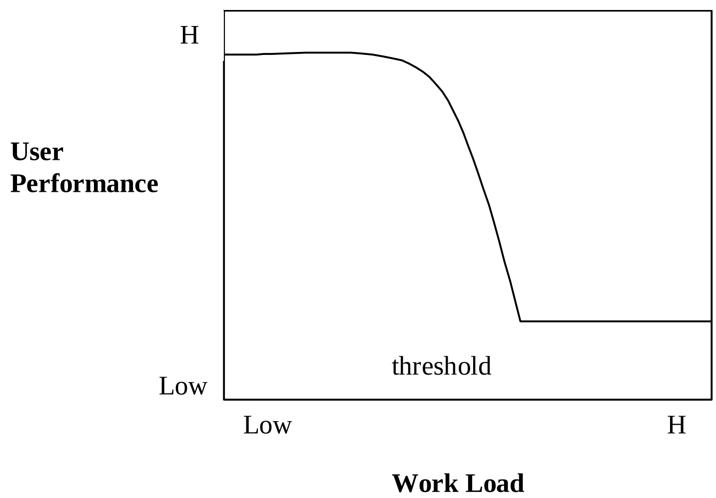


Figure 2.19. Expectation of user performance decrements with increases in workload.

The assessment of workload is a direct assessment of the class of difficulties that operators confront when performing an assigned task (Gopher and Donchin 1986). If disparity exists between task demands and a person's available resources, changes can be made to either the capacity of the person or the design of the task. However, the measurement of task difficulty itself is not straightforward (Gopher and Donchin 1986). Additionally, mental workload is not a directly observable trait of human-machine interactions because the operator easily masks it. A person has the ability to compensate for increases in task difficulty in order to maintain a particular level of task performance. Actions to incite changes in the mental workload associated with a task are coupled with influencers of attention and information processing capacity. Valid assessments of mental workload could inform the allocation of functional tasks between human and machine, comparisons of system designs (interfaces and tasks designs), the development of systems that can adapt task difficulty or function allocation in response to extreme levels (high and low) operator workload (Sanders and McCormick 1993).

Wierwille, Rahimi, and Casli evaluated 16 measures of mental workload to reveal 7 that were of suitable sensitivity for the domain of the pilot's task, the remaining 9 exhibiting high levels of variability (Wierwille, Rahimi and Casali 1985). Perhaps the most notable contribution of this work was the conclusion that researchers should never arbitrarily choose a workload measure.

The measurement of pupillary response is a physiological approach to the assessment of mental workload. O'Donnell and Eggemeier (1986) describe this class of workload metrics through sensitivity, diagniosticity intrusiveness, implementation and operator acceptance. Table 2.16 provides a summary of attributes of physiological measures. Physiological measures of workload cannot be used interchangeably, for they together are inclusive of a wide range of sensitivities, biases, and confounds.

Table 2.16. Summary of physiological measures of mental workload (O'Donnel & Eggemeier, 1986).

	Sensitivity	Diagnosticity	Constraints
Physiological Mental Workload Assessments	Can discriminate levels of capacity expenditure in non-overload situations. Assesses the relative potential among design options, tasks or operating conditions.	Level of diagnosticity is dependent on the technique employed. Brain-related potentials are highly diagnostic, whereas changes in pupil diameter have global indexing qualities.	Instrumentation for data collection restricts experimentation to controlled settings. Intrusion on task is not an issue, but interference may be an issue (due to invasive nature of data collection). Some methods may appear invasive and intimidating to users.

Pupil diameter was selected in the context of this test plan based on its potential as a global measure of operator mental workload. The assessment of pupil diameter, when appropriately applied, has the capacity to index combinations of resource utilization (O'Donnell and Eggemeier 1986). This aspect of the measure is appealing in a study to determine if multimodal feedback triggers significantly higher or lower demands upon a user’s information processing.

The impact of cognitive processing on pupillary response was realized over a hundred years ago when pupillary dilation was observed in response to stimuli and mathematical processing tasks (Beatty and Lucero-Wagoner 2000) The advent of serious, modern empirical work, however, did not increase until the

1960s with the publications of findings that pupillary activity was generated by mental activities (Andreassi 2000; Beatty and Lucero-Wagoner 2000).

Information processing is correlated to increases in pupil size above a baseline level and increase as processing increases (Boff and Lincoln 1988). Pupil diameter is said to reach maximum size when memory load is presumably highest (O'Donnell and Eggemeier 1986). Effectively, pupil dilation means increased bandwidth for information to be received to the photoreceptors on the retina (because a larger diameter means more light passes through), and sent to the brain via the optic nerve. This theory has been upheld by correlations between pupillary response to interest, arousal, information processing requirements and mental calculation (Andreassi 2000). Significant differences in pupil diameter are small in nature (on the order of .01 mm), but have been found both highly consistent and reliable. In fact, some have asserted the possibility of comparing relative workload across several different tasks (Beatty 1982). (O'Donnell and Eggemeier 1986) further state in their assessment of the state of the science in 1986, that pupil size will remain one of the most valuable indices of cognitive workload when properly used in the laboratory.

The analysis of pupillary response is transferable to several different applications, but within the confines of a laboratory setting. It is very useful in the assessment of different design options. The associated experimental tasks usually focus on information processes such as memory (e.g., digit span), language (e.g., grammatical, semantic categorization, sentence encoding, letter matching), reasoning (e.g., arithmetic), and perceptions (e.g., auditory, visual) (for a complete review, see Boff and Lincoln 1988).

Backs and Walrath (1992) conducted a two-part experiment with motivations similar to the present test plan. They applied performance measures, eye movement data, and pupillary response data to examine the effects of coding (color and symbolic) in visual displays of varied complexity. Pupillary response, in this study was sensitive to the information processing demands (based on complexity of task) and to physical display parameters (monochrome vs. color-coded display). In one experiment, greater dilation was found to occur during exhaustive search than during self-terminating search. Pupil diameter was found to be significantly larger to monochrome, low-density displays than to color-coded, low-density displays. Their experiment champions the efforts of combining ocular measures with 'traditional measures' or response and accuracy. The authors disclosed potential issues with the measurement of pupillary response.

Boff and Lincoln (1988) point out several constraints to this measure in the assessment of cognitive processing, which appear below. The impact of each issue on the current testing plan in terms of validity and reliability are identified, as are the actions to mitigate the effects.

Issue 1:

The pupil will automatically change diameter in response to ambient lighting, eye movements (including blinks), and emotional states (and medications). These changes are large enough to mask the effects of workload, and contaminate the measurement, taking away from the construct validity of pupil diameter.

Mitigating Actions:

The testing took place in visual laboratory darkrooms, where the ambient light was held constant. Participants also completed training and instruction in the testing environment prior to actual task completion to acclimate their pupils to the environment.

Backs and Walrath (Backs and Walrath 1992) introduce a useful methodology. They took several measurements during each trial, in order to correct for the subject’s initial pupil diameter, and is followed in the mining of the pupil data. The process is outlined:

- Assess baseline pupil diameter: pupil diameter at display screen onset
- Identify points of inflexion for constriction and dilation.
- Component A is a large constriction that peaks at about 950ms after the onset of the display
- Component B, followed Component A, and was characterized as a gradual dilation where peak latency varied with search time
- Subtract baseline pupil diameter from the two components, then measure the peak to peak different between component A and component B

Issue 2:

Analyses of pupil diameter cannot diagnose which *type* of resource is experiencing more demands than others. This is a deficiency in the construct validity of pupil diameter.

Mitigating Actions:

Conclusions are being made in terms of global indices of workload in the presence and absence of multimodal feedback; other constructs (in this case visual processing) are used to aid in the explanation of which resources are taxed in the human-machine system

Issue 3:

Near-vision pupillary reflex; the exposure to stimuli at close viewing distances over a long period of time can constrict the pupil. This could contaminate the data, and increase the deficiency of the measure.

Mitigating Action:

Changes in luminance were minimized between screens, (see Backs and Walrath 1992)

Issue 4:

Pupil diameter is highly variable and can fluctuate as much as 20% over a period of several seconds in static stimulus conditions. While the overall changes are small in magnitude (.01 mm), it is indicative of low sensitivity and reliability of the measure. Pupil diameter data tends to be ridden with noise.

Mitigating Actions:

1. Large sample sizes were gathered from each subject (60 Hertz sample rate);
2. The baseline methods introduced by Backs and Walrath (Backs and Walrath 1992), and outlined in Issue 1, were used in analyses; and
3. If variability in the data still prohibits a robust analysis, turn to more powerful analyses, utilizing the statistical concept of multifractality, and other time series analysis techniques (see Shi, Moloney, Emery et al. 2003).

Issue 5:

Studies of pupillary response are faced with the problem removing blink artifacts. A blink generally lasts about 70-100 msec. (producing an artifact spanning 4-6 observations under 60 Hz sampling) during which time the camera registers loss and a pupil diameter of zero is recorded. To complicate matters, there are times when ‘partial blinks’ are recorded before and after the zero value. This translates into unrealistically high and low values recorded by the eye tracker.

Mitigating Actions:

1. Utilize a blink-removal procedure where all zero values are removed, in addition to extreme values occurring within 6 observations on either side of the data.

2. The baseline procedure described in issue 1 will also deter this noise in the data stream.

In addition to pupillary response, eye movements, collected mainly to profile visual processing (as discussed in the subsequent section), will also be used to observe workload levels. An increase in saccadic movement has, in previous work, been found indicative of higher levels of workload (May, Kennedy et al. 1990). Additionally, high levels of workload are thought to decrease a person's ability to recognize cues in his or her peripheral vision (Rantanen and Goldberg 1999).

Visual Processing

A considerable number of studies concerning visual and cognitive processing have emerged in the past twenty years (Rayner 1998). Visual processing may be characterized through eye movements or psychophysiological measures such as EEGs or MRIs. This placed emphasis on extracting global indices of visual processing and attention through the record of eye movements. Compared to other psychophysiological methods, eye-tracking is minimally task-intrusive and typically agreeable by the participants. The eye-tracking field has made great advances in the past twenty years, mainly due to the development of more sophisticated tools and techniques for tracking eye movements (Rayner 1998). This section will detail the way eye movements are characterized, and the typical inferences drawn from these characterizations.

Eye movements can indicate a person's spatial focus of attention on a display (Goldberg and Kotval 1999). In terms of information processing, the movement of the eye is typically characterized by the attributes of saccades and fixations. A visual object is seen most clearly when it appears on the fovea, a small area in the center of the retina. A saccade is a ballistic, immediate movement of the eye that positions visual information on the central foveal region. Fixations are relatively still periods of 200-600 msec that occur between saccades (depending on the source) (Jacob 1991). Information is obtained primarily during fixations, and not saccades, a function of both physiological function and central processing (Rayner 1998). The saccades serve the role of placing specific images on the fovea, and the fixations are involved in the detailed visual analysis (Scott 1993).

Information is usually absorbed within a fixation in the central vision, while peripheral vision influences the choice of the next saccade (Scott 1993). Together, the central and peripheral vision work in tandem with cognition and memory to acquire visual information, and generated a scan path of saccadic eye movements and fixations. However, it is not entirely appropriate to infer that visual cue processing is occurring in conjunction with a fixation. Longer fixations are not always associated with greater extraction of information from the visual environment. It is often assumed that long fixations are more likely to occur in conjunction with visual information that is unfamiliar to the user, infrequent, or out of context (i.e., presenting higher levels of information). Wickens (1992) summarizes several studies where this is not always true. In one comparison of expert and novice users, the two groups spent roughly the same amount of time fixating. However, this was attributed to the fact that novice users were not aware of the information to be gathered, while and experts were highly efficient in visual search and information processing.

Theories of attention and cognitive control influence the scan path and general strategies for visual assessment. A person does not process the entire set of visual cues available in the environment; only those features that are most relevant and salient to current goals (Sanders and McCormick 1993). Additionally, not all saccades are visually guided; they may be reactive saccades in response to a sound, or saccades in anticipation, marked by voluntary movement to a location where information is expected to turn up. Knowledge of visual scan path does, however, lend itself to explanation of a person's internal mental models and expectations that drive selective attention (Wang and Stern 2001).

O'Donnell and Eggemeier (1986) state that "The absolute position of the eye at any point in time can be used to infer the information required to carry out a task, and many studies have used this type of measure to determine the processing requirements of a task," (O'Donnell and Eggemeier). Scan patterns can be an indication of global workload associated with the task. Table 2.17 provides a brief summary of how different domains apply eye-tracking techniques to answer a variety of research inquiries.

Table 2.17. Overview of domains applying eye tracking techniques.

Domain	Reference
Computer interface design evaluation technique (standard GUI interfaces)	(Goldberg and Kotval 1999)
Understanding human interaction with GUIs to better inform theories of human-computer interaction, and feed computational models of human-computer interaction	(Byrne 2001)
Assessing visual search of food nutrition labels. Analyzed scan patterns to determine relevant areas of interest	(Goldberg, Probart and Zak 1999)
Modeling focus of attention in collaborative environments	(Stiefelhagen, Yang and Waibel 1999)
Understanding the visual search patterns of users with visual impairments	(Jacko, Barreto, Scott et al. 2002)
Modeling eye movements while driving and changing lanes	(Salvucci and Liu 2002)
Modeling Situation awareness and mental workload in a aircraft combat situation	(Svensson, Angelborg-Thanderz, Sjöeberg and Olsson 1997)

In the context of this test plan, an account of a user’s visual/cognitive processes provided additional explanation for results that emerge from the performance data that were not altogether intuitive. Analyses of eye movements, including attributes of saccades and fixations, provided clarification on the mechanisms and resources used in the acquisition of visual information pursued during the course of interaction. This experiment will report fixation duration, saccade duration, and saccade to fixation ratio, measures of the efficiency and efficacy of visual information retrieval. These functions are driven by both the underlying physiological mechanisms of the visual sensory system, but also a function of working memory and cognition.

Goldberg and Kotval (1999) created a taxonomy of eye movement measurements. This taxonomy is comprised of measures of visual search, and more importantly to this test plan, measures of visual processing. All measures fall into one of these classifications, and are further clarified by what part of the dynamic environment they can explain: temporal and/or spatial. The authors categorize the following variables as measures of information processing: number of fixations, fixation duration, and fixation/saccade ratio. These measurements of the depth and breadth of visual processing lend value in bringing to the surface some of the covert processes of situational awareness and information processing (See Backs and Walrath 1992).

Reliability and Validity of Eye tracking measures

The utilization of eye tracking equipment to capture movement and pupillary response often proves challenging when working with a population that deviates from the norm, like the aging or visually impaired. In some instances, the ability of the optics to capture an accurate image hinges on a delicate balance between the participant (i.e. the shape of their head or eye, clouded lens, floaters) and environmental factors (ambient lighting, viewing angle, etc.). In addition, in consideration of the propensity for participants with AMD to employ E-centric viewing strategies, the ability of the eye tracker to accurately pinpoint the location of fixation and saccades is not reliable. However, through the capture of the pupil and corneal reflection and their movement, the eye tracking controller can still aptly detect fixations, inter-fixations (saccades), and their respective durations. Because the number of fixations actually captured by the eye tracking system could deviate between participants, any measures based on fixation or saccade frequency were excluded. Instead alternative summary measures were generated based on the duration of each fixation and saccade. These included: saccade duration, fixation duration and saccade to fixation duration ratio.

Equipment issues:

While the validity of eye tracking techniques is sufficient, concerns arise with the reliability of the measures. Typically, eye trackers can estimate point of gaze within 1 to 2 degrees visual angle. In response to the sensitivity and repeatability of the measurements themselves, because this study does not considered the physical scan patterns and distribution as they relate to specific regions on the display, the sensitivity of the equipment with respect to point of gaze is not an issue. However, error may be introduced if the equipment is not reliably capturing fixations; signals are susceptible to environmental light, facial features, and differences between subjects. A large enough sample size will be taken from each subject to minimize overall variability in signal (60 Hz). Additionally, the equipment is calibrated to the movement of each subject's eye at the beginning of an experimental trial.

Data Inference Issues:

"Assuming a researcher, interested in studying the usability of a human-computer interface, is not scared off by the technical and data extraction problems . . .there is still the issue of making sense out of eye tracking data. How does the usability researcher relate fixation patterns to task-related cognitive activity?" (Jacob and Karn 2003)

It is therefore important to note how a fixations and saccades were derived from the raw eye stream data. For the purposes of this study, fixations were identified from the data based on continuous point of gaze within a 1x1 degree of visual angle on the display, for at least 100 ms. Saccades were identified as the period between the fixations.

Another issue in terms of both psychometric reliability and validity is that a fixation is not always indicative of visual processing. The duration of a fixation is not wholly monotonically related to quantity of information processing. A fixation does not necessarily mean that a person processed what was within their visual point of gaze (or that they even perceived it) (Pew 2000), but can indicate that information processing occurred.. That said, the eye tracking metrics collected will be regarded as descriptive account of the visual processes a person undertakes during this task and the extent of demands on the visual sensory channel in corroboration with cognition.

2.5. Detailed Hypotheses

In light of the experiment and the several dependent variables collected in this investigation, the original hypotheses, introduced in Chapter 1, are further clarified in this section. Chapter 3, Chapter 4 and Chapter 5 present the results, analyses, and the results that emerged from of the handheld PC and desktop PC experiments, and their comparison, respectively. At the final section of each chapter the relevant hypotheses are considered with supporting evidence and conclusions.

Hypothesis 1

For all users, there is a point of diminishing return for performance gains attributed to increases in icon size, set size, spacing, and overall screen density.

a. The point of diminishing return is dependent on a user's visual capacity, including clinical measures of vision, visual attention, and subjective visual functioning.

b. Set size, icon size and spacing will influence the components of interaction (e.g., visual search, icon acquisition, dragging and dropping) in different ways and magnitudes, also dependent on the visual capacity of the participant.

c. The negative effects of extremely dense (many icons with small inter-icon spacing) interfaces will be amplified on the desktop PC more than the handheld PC, due to the demands required to scan the larger interface, normalized based on display size.

d. Independent of performance, the demands placed upon the users' covert processes, such as visual processing and cognitive workload will also demonstrate a point diminishing return in the presence of changes to the visual interface.

Hypothesis 2

The potentially positive influence of auditory feedback on the drag and drop task is effectively masked by the complexity of the task (multiple icons and multiple targets as compared to the single file – single folder task used in previous studies).

a. The presence of auditory feedback will not affect the visual search component of this task

b. As the number of icons and potential drop targets increase, the presence of auditory feedback can have detrimental effects on the interaction.

Hypothesis 3

Measures of ocular health and visual function are predictors of performance in the required task.

a. Certain components of the interaction (e.g., the visual search, drag and drop) are more susceptible to the negative impacts of limited ocular functioning.

b. The components of the interaction are influenced by each measure of ocular health and visual function to a different degree.

c. Different measures of ocular ability can delineate which components of the task will be executed in a less efficient manner, following the speed-accuracy tradeoff common to most HCI tasks.

d. The predictive power of attributes of visual function used to classify the outcome of interactions will differ in terms of the amount of influence on the task.

e. The influence of ocular functioning on task interaction greatly overrides other normal, age-related declines in mental and physical health.

f. Non-clinically-acquired measures of visual function, such as users' perception of the impact of visual dysfunction on their activities of daily living (VFQ), and functional visual attention are more powerful predictors of the outcomes of the interaction than clinical factors.

Hypothesis 4

The interactions of those users in the handheld PC group will be less efficient than those users in the desktop PC group.

a. The motor skill required by the input device will cause users to slow their performance at different points in the interaction, demonstrating a speed–accuracy tradeoff not readily observable in the desktop PC condition.

b. Users with visual impairments will experience more performance decrements with respect to interactions on a handheld device than the visually healthy cohorts.

c. The most detrimental component of the handheld PC to interactions for individuals with visual impairments is the size of the icons.

d. Normal age-related declines to mental and physical health will be amplified by the interaction style required by the handheld PC.

This study is best classified as an exploratory study. That said, the analyses of exploratory studies are not typically straightforward. This dissertation takes advantage of modeling techniques such as linear regression and logistic regression to understand the relative impact of the various factors on the dependent measures. In other words, the study aims to establish the subset of factors that are optimal in the characterization this interaction: clinically acquired measures of visual function and health, non-clinical measures of visual function and health, personal characteristics or interface augmentations. With this knowledge, future work can target the further exploration of actions to mitigate the impact of visual impairment through strategic design.

2.6 REFERENCES

Andreassi, J. L. (2000). Psychophysiology: Human behavior & physiological response. Mahwah, NJ, Lawrence Erlbaum Associates.

Arditi, A. and R. Cagenello (1993). "On the statistical reliability of letter-chart visual acuity measurements." Investigative Ophthalmology and Visual Science **34**(1): 120-129.

Backs, R. W. and L. C. Walrath (1992). "Eye movement and pupillary response indices of mental workload during visual search of symbolic displays." Applied Ergonomics **23**: 243-254.

Beatty, J. (1982). "Task-evoked pupillary responses, processing load, and the structure of processing resources." Psychological Bulletin **91**: 377-381.

Beatty, J. and B. Lucero-Wagoner (2000). The pupillary system. Handbook of psychophysiology. J. T. Cacioppo, L. G. Tassinary and G. G. Berntson. Cambridge, UK, Cambridge University Press: 142-162.

Bird, A. C., N. M. Bressler, S. B. Bressler, I. H. Chisolm, G. Coscas, M. D. Davis, P. T. V. M. d. Jong, C. C. W. Klaver, B. E. K. Klein, R. Klein, P. Mitchell, J. P. Sarks, S. H. Sarks, G. Soubrane, H. R. Taylor and J. R. Vingerling (1995). "An international classification and grading system for age-related maculopathy and age-related macular degeneration: The international ARM epidemiological study group." Survey of Ophthalmology **39**(5): 367-374.

Boff, K. R. and J. E. Lincoln (1988). Engineering data compendium: Human perception and performance. New York, NY, John Wiley and Sons.

Bressler, N. M., S. B. Bressler, S. K. West, S. L. Fine and H. R. Taylor (1988). "The grading and prevalence of macular degeneration in Chesapeake Bay waterman." Archives of Ophthalmology **107**(June): 847-852.

Bressler, N. M., S. B. Bressler, S. K. West, S. L. Fine and H. R. Taylor (1989). "The grading and prevalence of macular degeneration in Chesapeake Bay Waterman." Archives of Opthalmology **107**: 847-852.

Brewster, S. A. (2002). "Overcoming the Lack of Screen Space on Mobile Computers." Personal and Ubiquitous Computing **6**(3): 188-205.

Byrne, M. D. (1993). Using icons to find documents: Simplicity is critical. Proceedings of the SIGCHI conference on Human factors in computing systems, Amsterdam, The Netherlands.

Byrne, M. D. (2001). "ACT-R/PM and menu selection: applying a cognitive architecture to HCI." International Journal of Human Computer Studies **2001**(55): 41-84.

Coyle, J. T. (2003). "Use it or lose it - do effortful mental activities protect against dementia?" New England Journal of Medicine **348**: 2489--2490.

Dain, S. J. (2004). "Clinical colour vision tests." Clinical and Experimental Optometry **87**(4-5): 276-293.

Derrick, W. L. (1988). "Dimensions of operator workload." Human Factors **30**(1): 95-110.

Edwards, P. J., L. Barnard, V. K. Emery, J. S. Yi, K. P. Moloney, T. Kongnakorn, J. A. Jacko, F. Sainfort, P. R. Oliver, J. Pizzimenti, A. Bade, G. Fecho and J. Shallo-Hoffmann (2004). Strategic Design for Users with Diabetic Retinopathy:
Factors Influencing Performance in a Menu-Selection Task. Sixth International ACM SIGCAPH Conference on Assistive Technologies (ASSETS 2004), Atlanta.

Emery, V. K., P. J. Edwards, J. A. Jacko, K. P. Moloney, L. Barnard, T. Kongnakorn, F. Sainfort and I. U. Scott (2003). Towards achieving universal usability for older adults through multimodal feedback. ACM Conference on Universal Usability (CUU), Vancouver, British Columbia, Canada.

Emery, V. K., J. A. Jacko, T. Kongnakorn, V. Kuruchittham, S. Landry, G. M. N. III, A. Sears and J. Whittle (2001). Identifying critical interaction scenarios for innovative user modeling. 10th International Conference on Human-Computer Interaction, New Orleans, LA.

Everett, S. P. and M. D. Byrne (2004). "Unintented effects: Varying icon spacing changes users' visual search strategy." CHI Letters **6**(1): 695-702.

Goldberg, J. H. and X. P. Kotval (1999). "Computer interface evaluation using eye movements: methods and constructs." International Journal of Industrial Ergonomics **24**(6): 631-645.

Goldberg, J. H., C. K. Probart and R. E. Zak (1999). "Visual Search of Food Nutrition Labels." Human Factors **41**(3): 425-437.

Gopher, D. and E. Donchin (1986). Workload- an examination of the concept. Handbook of Perception and Human Performance. L. K. K. R. Boff and J. P. Thomas. New York, NY, John Wiley and Sons: 41-49.

International Council of Ophthalmology. (2005). "International Standards: Visual Standards - Aspects and Ranges of Vision Loss." Retrieved May 11, 2005, from www.icoph.org/standards/visionres.html#c.

Jacko, J. A., L. Barnard, T. Kongnakorn, K. P. Moloney, P. J. Edwards, V. K. Emery and F. Sainfort (2004). "Isolating the effects of visual impairment: Exploring the effect of AMD on the utility of multimodal feedback." CHI Letters **6**(1): 311-318.

Jacko, J. A., A. B. Barreto, J. Y. M. Chu, H. S. Bautsch, G. J. Marmet, I. U. Scott and R. H. Rosa (2000). Using eye tracking to investigate graphical elements for fully sighted and low vision users. ACM conference on Eye Tracking Research Applications (ETRA), Palm Beach, FL.

Jacko, J. A., A. B. Barreto, I. U. Scott, J. Y. M. Chu, H. S. Vitense, F. T. Conway and W. B. Fain (2002). "Macular Degeneration and visual icon use: Deriving guidelines for improved access." Universal Access in the Information Society **1**(3): 197-206.

Jacko, J. A., V. K. Emery, P. J. Edwards, M. Ashok, L. Barnard, T. Kongnakorn, K. P. Moloney, F. Sainfort and I. U. Scott (2004). "The effects of multimodal feedback on older adults' task performance given varying levels of computer experience." Behaviour & Information Technology **23**(4): 247-264.

Jacko, J. A., K. P. Moloney, T. Kongnakorn, L. Barnard, P. J. Edwards, V. K. Emery, F. Sainfort and I. U. Scott (2005). "Multimodal feedback as a solution to ocular disease-based user performance decrements in the absence of functional visual loss." International Journal of Human-Computer Interaction. **18**(1).

Jacko, J. A., R. H. Rosa, I. U. Scott, C. J. Pappas and M. A. Dixon (1999). Linking visual capabilities of partially sighted computer users to psychomotor task performance. 9th International Conference on Human-Computer Interaction, Munich, Germany.

Jacko, J. A., I. U. Scott, A. B. Barreto, H. S. Bautsch, J. Y. M. Chu and W. B. Fain (2001). Iconic visual search strategies: A comparison of computer users with AMD versus computer users with normal vision. 10th International Conference on Human-Computer Interaction, New Orleans, LA.

Jacko, J. A., I. U. Scott, F. Sainfort, L. Barnard, P. J. Edwards, V. K. Emery, T. Kongnakorn, K. P. Moloney and B. Zorich (2003). "Older adults and visual impairment: What do exposure times and accuracy tell us about performance gains associated with multimodal feedback?" CHI Letters **5**(1): 33-40.

Jacko, J. A., I. U. Scott, F. Sainfort, K. P. Moloney, T. Kongnakorn, B. Zorich and V. K. Emery (2002). Effects of multimodal feedback on the performance of older adults with normal and impaired vision. Lecture notes in computer science. C. S. Noelle Carbonell. NEED CITY, Springer. **2615:** 3-22.

Jacob, R. J. K. (1991). "The use of eye movements in human-computer-interaction techniques: what you look at is what you get." ACM Transactions on Information Systems **9**(3): 152-169.

Jacob, R. J. K. and K. S. Karn (2003). Eye Tracking in Human-Computer Interaction and Usability Research: Ready to Deliver the Promises (Section Commentary). The Mind's Eye: Cognitive and Applied Aspects of Eye Movement Research. J. Hyona, R. Radach and H. Deubel. Amsterdam, Elsevier Science: 573-605.

Kahneman, D. (1973). Attention and effort. Englewood Cliffs, NJ, Prentice-Hall.

Leonard, V. K., P. J. Edwards and J. A. Jacko (2005). Informing accessible design through self-reported quality of visual health. Human Factors and Ergonomics Society, Orlando.

Mangione, C. M., S. Berry, K. Spritzer, N. K. Janz, R. Klein, C. Owsley and P. P. Lee (1988). "Identifying the content area for the 51-item National Eye Institute Visual Function Questionnaire: Results from focus groups with visually impaired persons." Archives of Ophthalmology **116**: 227-233.

Mangione, C. M., P. P. Lee, P. R. Gutierrez, K. Spritzer, S. Berry and R. D. Hays (2001). "Development of the 25-Item National Eye Institute Visual Function Questionnaire." Archives of Ophthalmology **119**: 1050-1058.

May, J. G., R. S. Kennedy, M. C. Williams, W. P. Dunlap and J. R. Brannan (1990). "Eye movement indices of mental workload." Acta Psychologica **75**: 75-89.

Microsoft Developers Network. (2004). "XXX." from www.msdn.com.

O'Donnell, R., & and F. T. Eggemeier (1986). Workload Assessment Methodology. Handbook of Perception and Human Performance. K. R. Boff, L. Kaufman and J. P. Thomas. New York, NY, John Wiley and Sons**:** 42-1-42-49.

Ormerod, L. D. and S. Mussatt. (2005). "Initial low vision assessment." Retrieved May 11, 2005, from http://www.vhct.org/case3002/test_visual_ac.shtml.

Pascoe, J., N. Ryan and D. Morse (1999). Issues in Developing Context-Aware Computing. Handheld and Ubiquitous Computing, Karlsruhe, Germany, Berlin; New York: Springer, c1999.

Pelli, D. G., J. G. Robson and A. J. Wilkins (1988). "The design of a new letter chart for measuring contrast sensitivity." Clinical Visual Science **2**: 187-199.

Pew, R. W. (2000). The state of situation awareness measurement: heading toward the next century. Situation Awareness Analysis and Measurement. M. Endsley and D. Garland. Mawah, NJ, Lawrence Erlbaum Associates**:** 33-47.

Rantanen, E. M. and J. H. Goldberg (1999). "The effect of mental workload on the visual field size and shape." Ergonomics **42**(6): 816-834.

Rayner, K. (1998). "Eye movements in reading and information processing: 20 years of research." Psychological Bulletin **124**(3): 372-422.

Salvucci, D. D. and A. Liu (2002). "The time course of a lane change: Driver control and eye-movement behavior." Transportation Research **F**(5): 123-132.

Sanders, M. S. and E. J. McCormick (1993). Human Factors in Engineering and Design. New York, McGraw-Hill.

Scott, D. (1993). Visual search in modern human-computer-interaction. Visual Search 2. A. Gale, K. Carr and D. Brogan. New York, Taylor & Francis**:** 175-189.

Scott, I. U., W. J. Feuer and J. A. Jacko (2002). "Impact of graphic user interface screen features on computer task accuracy and speed in a cohort of patients with Age-related Macular Degeneration." American Journal of Ophthalmology **134**(6): 857-862.

Sears, A., J. A. Jacko, B. Brewer and L. Robelo (1998). A framework for visual icon design. 42nd Annual Meeting of the Human Facotrs Societ, Santa Monica, CA, Human Factors Society.

Sheridan, T. B. (2002). Humans and automation: System design and research issues. New York, NY, John Wiley and Sons.

Shi, B., K. Moloney, V. K. Emery, J. Jacko, F. Sainfort and B. Vidakovic (2003). Multifractal Discrimination Model of High-Frequency Pupil-diameter Measurements. INFORMS, Atlanta.

Stiefelhagen, R., J. Yang and A. Waibel (1999). Modeling focus of attention for meeting indexing. Seventh ACM international conference on Multimedia, Orlando, Florida, ACM Press.

Svensson, E., M. Angelborg-Thanderz, L. Sjöeberg and S. Olsson (1997). "Information complexity: mental workload and performance in combat aircraft." Ergonomics **40**(362-380).

Tiffin, J. and E. J. Asher (1948). "The Purdue Pegboard: Norms and studies of reliability and validity." Journal of Applied Psychology **32**: 234-247.

Trenerry, M. R., B. Crosson, J. DeBoe and W. R. Leber (1990). VSAT: Visual Search and Attention and Attention Test. Lutz, FL, PAR: Psychological Assessment Resrouces, Inc.

Tsang, P., & Wilson, G. F. (1997). Mental Workload. Handbook of Human Factors and Ergonomic. G. Salvendy. New York; NY, John Wiley and Sons**:** 417-449.

Uvijls, A., B. P. Leroy, M. Leys, A. D. Rouck and P. Kestelyn (1998). "A large field Panel D-15 test for low vision patients." Bull Soc Belge Ophtalmol **1998**(270): 57-65.

Vitense, H. S., J. A. Jacko and V. K. Emery (2002). "Foundation for improved interaction by individuals with visual impairments through multimodal feedback." Universal Access and the Information Society **2002**(2): 76-87.

Vitense, H. S., J. A. Jacko and V. K. Emery (2003). "Multimodal feedback: an assessment of performance and mental workload." Ergonomics **46**(1-3): 68-87.

Wang, L. and J. Stern (2001). "Saccade initiation and accuracy in gaze shifts are affected by visual stimulus significance." Psychophysiology **38**(20001): 64-75.

Ware, J. E., M. Kosinski and S. D. Keller (1995). How to Score the SF-12 Physical and Mental Health Summary Scales. Boston, MA, The Health Institute, New England Medical Center.

Wickens, C. (1992). Attention in Perception and Display Space. Engineering Psychology and Human Performance. New York, HarperCollins**:** 74-115.

Wickens, C. D. (1980). The structure of attentional resources. Attention and performance. R. Nickerson. Englewood Cliffs, NJ, Erlbaum. **VIII**.

Wickens, C. D. and Y. Lui] (1988). "Codes and modalities in multiple resources: A success and qualification." Human Factors, 30(5), 599-616. **30**(5): 599-616.

Wickens, C. D. and Y.-Y. Yeh (1982). The dissociation of Subjective Ratings and Performance. EEE International Conference on Cybernetics, Seattle, WA.

Wierwille, L., S. Tijerina, T. Kiger, E. Rockwell, E. Lauber and A. Bittner, Jr. (1996). Heavy vehicle driver workload assessment. Task 4: Review of workload and related research. Washington D.C., U.S. Department of Transportation, National Highway Traffic Safety Administration.**:** 1-96.

Wierwille, W. W., M. Rahimi and J. g. Casali (1985). "Evaluation of 16 measures of mental workload using a simulated flight task emphasizing mediation activity." Human Factors **27**(5): 489-502.

Wilson, G. F. (2001). "An analysis of mental workload in pilots during flight using multiple psychophysiological measures." The International Journal of Aviation Psychology **12**(1): 3-18.

CHAPTER 3

HANDHELD PC:
ANALYSES, RESULTS & CONCLUSIONS

3.1 <u>Introduction</u>

This chapter presents the analyses, results and conclusions associated with the different facets of participant interaction with the handheld computer, on the tasks and variables detailed in Chapter 2. As outlined in Chapter 2, thirteen volunteers from the NSU College of Optometry patient pool and associates of NSU staff participated in the handheld component of this study. Ten participants' only ocular diagnosis was some level of AMD, and the remaining three participants were visually healthy, age-matched controls with no ocular disease present. The demographic backgrounds and ocular profiles of these individuals were summarized in Chapter 2, along with the experimental procedures.

Overall, the thirteen participants performed quite well in the execution of the task. Twelve participants completed all 162 trials required by the task, and one participant completed 92 (57.4% of the trials) (this participant's session was terminated because of fatigue). On average, participant task time on the handheld lasted 19.71 min (standard deviation = 12.69 min; median = 15.45 min).

In terms of overall trial accuracy, the final icon dropped into the pile was correct on 95.7% of the trials across all participants. The high rates of task accuracy demonstrate the true potential for handheld computers to be effectively used by both the general aging population and those who are aging with visual impairments (such as those common to AMD). However, in consideration of the high rate of task accuracy, these overall summary outcome measures of task performance hold little value in the assessment of the impact of the numerous predictor variables. For this reason, the measures of the various phases of implicit and explicit interaction were of particular interest in providing a more illustrative account of task interaction.

As this investigation was largely an exploratory effort, regression (linear and logistic) posed notable potential in satisfying the established hypothesis. That is, regression analyses provided the potential to ascertain which factors were most influential on handheld interaction for this population; which factors were fundamentally driving the different components of the interaction and the quantification of these effects in relation to the outcomes and each other.

While analyses using group comparisons have more commonly been employed in previous research involving HCI and visual impairments, the regression approach was taken with this data set for several reasons (Jacko, Barreto, Scott et al. 2002; Jacko, Barnard, Kongnakorn et al. 2004; Jacko, Moloney, Kongnakorn, Barnard, Edwards, Leonard et al. 2005). First, when participants in the current study were classified into groups based on visual acuity (as in the previous studies), the groups varied, in a non-uniform manner on other aspects of visual function, severity of AMD, and age. Second, while visual acuity can be indicative of AMD, several clinically-based ophthalmic studies have deliberately excluded visual acuity in the grading AMD severity (Bird, Bressler, Bressler et al. 1995) In more practical terms concerning functional vision, visual acuity is not the sole factor influencing how the individual perceives the computer interface. Also, Jacko and colleagues attributed several visual factors to drive performance during visual search for an icon, including visual acuity, contrast sensitivity, and visual field). (Jacko, Rosa, Scott, Pappas and Dixon 1999; Jacko 2000)

In consideration of the great number of measures taken to profile the various phases of the interaction, both implicit and explicit, regression provides the more efficient means by which to compare the relative influence of the various independent variables in relationship to each other. Regression enables the exploration of the impact on the components of interaction as well as the relationships existent among the predictors (Field 2000). The utility of regression in explaining computer interactions and visual ability was demonstrated by Edwards and colleagues (Edwards, Barnard, Emery et al. 2004; Edwards, Barnard, Emery et al. 2005). In this work regression in an appraisal performance variability for users with Diabetic Retinopathy with a drop-down menu under and various interface conditions. Prior to this Diabetic Retinopathy study, Scott Jacko and Feuer applied regression modeling to their examination of the factors affecting icon recognition and selection (Scott, Feuer and Jacko 2002, 2002).

The analyses and results presented in this chapter are divided into three sections according to the various dependent measures considered: 1) Efficiency, 2) Accuracy, and 3) Information Processing. Each dependent measure is considered in light of clinically-acquired visual factors, non-clinically-acquired visual factors, personal characteristics and interface features (i.e.. set size, spacing, and auditory feedback), and other extraneous task-related factors (e.g., location of target icons, and order effects). Finally, the qualitative data gathered from the participants in exit surveys are summarized and related back to the quantifiable assessments. This chapter concludes with a section linking the results back to those hypotheses relevant to the handheld computer task.

Tables 3.1 through 3.3 summarize the predictor variables considered in modeling the variance of the dependent variables, grouped according to *task-related factors, general participant-related factors*, *clinically-acquired ocular factors,* and *non-clinically-acquired ocular factors*. The VFQ measures are further summarized by Figure 3.1, which details the responses on all thirteen subscales. Because of the lack of diagnostic specificity in the individual VFQ subscales for this population sampled, only the VFQ Overall Score was employed in the statistical analyses. Furthermore, the inclusion of all subscales, which are so highly correlated, would have otherwise compromised the integrity of the analyses. Also included in these summary tables is a description of the levels observed for each predictor in the handheld participant group. In addition to those variables listed, statistical interactions were considered for AMD Score, VSAT and VFQ Overall Score with the controlled interface factors (SS, AF, and ISp).

Table 3.1. General participant-related factors.

Predictor Variable	Description	Observed Levels
Age	Participant's age at the time of the study	Range: 53 – 82 years Mean: 68.69 Median: 70
SF-12 Physical Component Score *SF-12 PCS*	Survey of participant's self reported physical health at the time of the experiment, from 1 (worst) to 100 (best)	Range: 28.64 - 60.46 Mean: 46.15 Median: 45.22
SF-12 Mental Component Score *SF-12 MCS*	Survey of participant's self-reported mental health at the time of the experiment, from 1 (worst) to 100 (best)	Range: 26.39 - 60.79 Mean: 46.74 Median: 48.61
Purdue Pegboard Test of Manual Dexterity *Dexterity*	Count of the average number of pins inserted into small holes in a board over three, 30 second trials, from 0 (worst) to 30 (best)	Range: 4.67-16.33 Mean: 11.49 Median: 12.33

Table 3.2. Clinically-acquired ocular factors.

Predictor Variable	**Description**	**Observed Levels**
LogMar Near Visual Acuity† *NVA*	Ability to visually focus on fine details at a distance of 40 cm, translated from Snellen acuity (e.g. 20/20). Scores can range from .1 (best) to 1(worst)	Range: 0.00-1.00 Mean: .7127 Median: .8000
Contrast Sensitivity† CS	Measure of how visible an image is before it becomes indistinguishable from a uniform field, from 0 (low) to 60 (high)	Range: 26.00 - 40.50 Mean: 33.5000 Median: 34.5000
AMD Severity Score† *AMD Score*	A diagnosis of severity of disease based on examination of photographs of the eye from no disease (0) to severe (4)	Range: 0-4.00 Mean: 1.1731 Median: 1.0000

†For NVA, CS and AMD Severity Score, weighted average of the best and worst eye (.75 * best + .25 * worst) approximated binocular visual field.

Table 3.3. Non-clinically-acquired ocular factors.

Predictor Variable	Description	Observed Levels
Visual Attention Test *VSAT*	A paper based assessment of sustained visual attention. Scores on the VSAT have been compared to an age-related normative sample, and recoded according to their relative percentile (based on age). The higher the percentile, the less severe the detected impairment. -At or below 2nd percentiles: Significant, impaired vision -At or between 3rd to 16th percentile: Suggestive, borderline impairment -At or above 17th percentile: Within the normal range of performance	Range: 1st -83rd percentiles Mean: 34.07 Median: 32.01 Incidence of categorical classification: Significant: 2 Borderline: 3 Normal: 8
NEI Visual Functioning Questionnaire-25 *VFQ*	Self-perceived assessment of visual function and daily activity, based on responses to the verbally administered NEI-VFQ-25. Scores are generated for each of the 13 subscales, and 1 overall VFQ score is calculated. Scores can range from 0 (maximum interference with daily functioning, or worse perceived visual function) to 100 (no interference, best possible perceived visual function).	Overall VFQ Score Range: 47.5-99.03 Mean: 82.83 Median: 88.20 (95%CI are provided for Overall Score and the other 12 subscales in Figure 3.1)

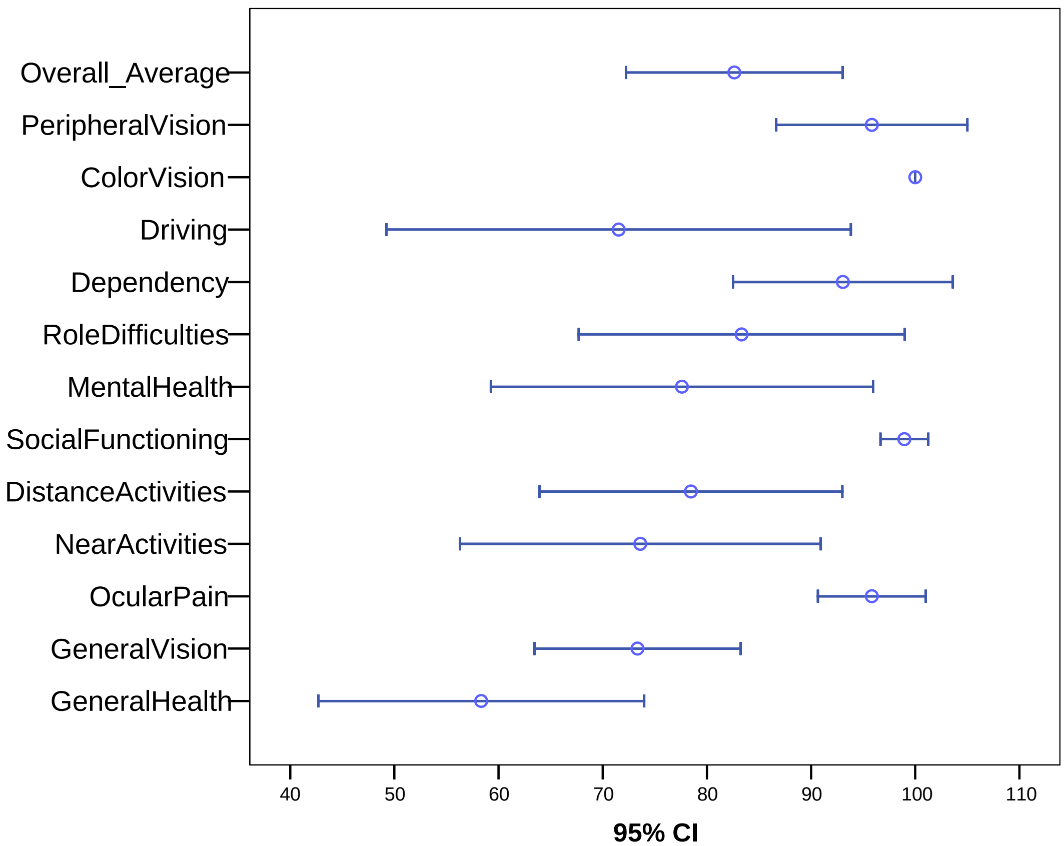


Figure 3.1. 95% Confidence intervals for the 13 VFQ subscales and overall average.

Table 3.4. Task-related factors.

Predictor Variable	Description	Observed Levels
Set Size (SSI)	The number of card icons presented for each trial	1: 4 card icons 2: 8 icons 3: 12 icons
Inter-Icon Spacing (ISp)	The number space between the card icons and drop piles (above and below)	1: ¼ icon 2: ½ icon 3: 1 icon
Auditory Feedback (AF)	Supplemental auditory feedback to communicate the position of the card for an accurate drop	0: AF absent 1: AF present
Column Location (Column)	The column where the target card icon is located for each trial	1: leftmost 2: middle 3: rightmost
Row Location (Row)	The row where the target card icon is located for each trial	1: top 2: 2nd from top 3: 2nd from bottom 4: bottom
Drop Location	The row number of where the correct drop pile for each trial was located	1: top 2: 2nd from top 3: 2nd from bottom 4: bottom
Trial Number (Trial #)	Sequential position of the trial within a participant's overall experimental session	Range: 0 -161

In the following sections, the specific analyses executed and the results are detailed, in three parts, based on the taxonomy of dependent variables: *Efficiency, Accuracy, and Information Processing*. Each set of dependent measures is modeled using clinically-acquired ocular factors and non-clinically-acquired ocular factors separately, while each model consistently considers general participant related factors (e.g., age) and task related factors (e.g., spacing, set size, trial number, given in Table 3.4). The predictor variables considered in the clinically-acquired ocular analyses were consistent between regressions on each outcome measure, as were those considered between the non-clinically-acquired prediction models. This was done to ensure that the relative effects between different task phases could be compared. A complete aggregated list of predictor variables follows:

Ocular Factors	Personal Factors	Interface Factors	Interaction Terms
Clinical	SF-12 PCS	SS	AMD Score * SS
NVA	SF-12 MCS	ISp	AMD Score * ISp
CS	Dexterity	AF	AMD Score * AF
AMD Score	Age	Column Location	VFQ * SS
Non-Clinical		Row Location	VFQ * ISp
VSAT Percentile		Drop Location	VFQ * AF
Overall VFQ Score		Trial Number	VSAT * SS
			VSAT * ISp
			VSAT * AF

3.2 Efficiency Measures

General Summary

Table 3.5 provides a synopsis of the efficiency outcome measures, and the distribution observed for each from this population's interaction with the handheld task. Figures 3.2 and 3.3 summarize the nature of the task through graphs of the means and standard error for each of the time-based efficiency measures. It is observed in Figure 3.2 that relative to TT, the majority of participants' time in each trial was spent in VS, then MT. This helps to prioritize potential design interventions, and substantiates the supposition that that the experimental task was visually demanding.

While TTHT and FTHT, as shown in Figure 3.2, are much smaller in scale compared to VST and MT, both are considered in the following analyses. The measures provide the direct means by which to assess the impact of the auditory feedback on the drop portion of the task. Despite the limited relative magnitude of the target highlight times, these measures are in fact components of the movement time measure. Moreover, because they were reported in several previous studies concerning the efficacy of feedback they afford comparisons to the prior findings (Jacko, Barreto et al. 2002; Vitense, Jacko and Emery 2002; Jacko, Scott, Sainfort et al. 2003; Vitense, Jacko and Emery 2003; Jacko, Moloney, Kongnakorn, Barnard, Edwards, Leonard et al. 2005). Figure 3.3 presents an overview of those efficiency measures based on the number of pixels over which the stylus travels with the icon attached. Through this graphic, the relative impact of the MV and ME on the overall stylus movement (DD) is assessed.

Table 3.5. Summary of efficiency measures of handheld interaction.

Efficiency Measures	
Trial Time-*TT* (msec)	Mean (Std. Error): 3959.85 (91.30) Median: 3001.50 Minimum: 1150.00 Maximum: 60362.00
Visual search time-*VST* (msec)	Mean (Std. Error): 2546.52 (89.50) Median: 1674.500 Minimum: 431.00 Maximum: 56250.00
Movement Time-*MT* (msec)	Mean (Std. Error): 1413.33 (20.56) Median: 1203.50 Minimum: 72.00 Maximum: 15880.00
Final target highlight time-*FTHT* (msec)	Mean (Std. Error): 5.70 (1.77) Median: 3.00 Minimum: 2.00 Maximum: 3026.00
Dragging Distance-*DD* (pixels)	Mean (Std. Error): 124.62 (2.12) Median: 106.98 Minimum: 18.71 Maximum: 1467.71
Movement Variability-*MV* (pixels)	Mean (Std. Error): 6.05 Median: .102 Minimum: .272 Maximum: 34.2
Movement Error-*ME* (pixels)	Mean (Std. Error): 10.46 (.177) Median: 8.58 Minimum: .387 Maximum: 68.05

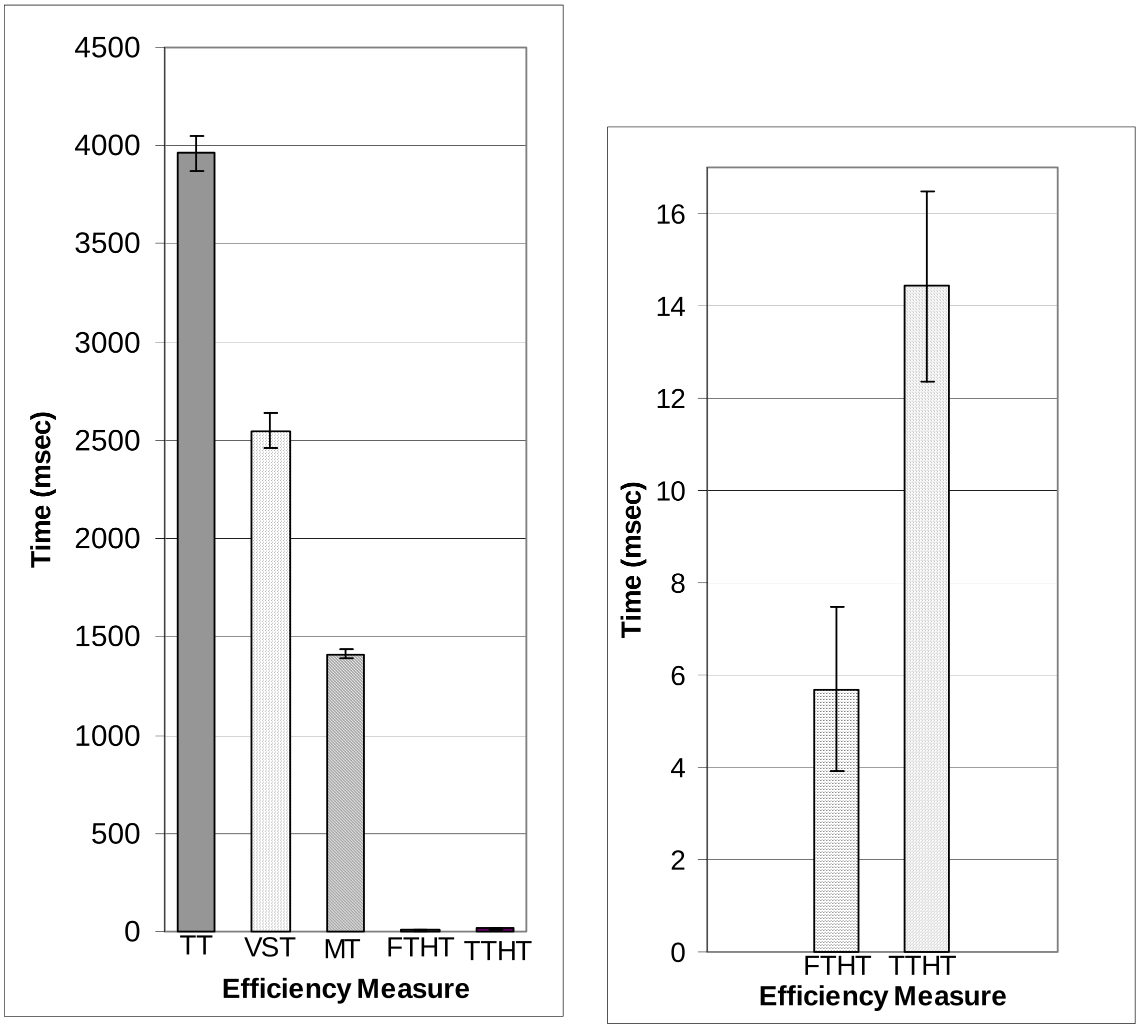


Figure 3.2. Summary of means and standard error for time-based efficiency scores all time-based measures; Means for FTHT and TTHT, unquantifiable in the overall efficiency graph.

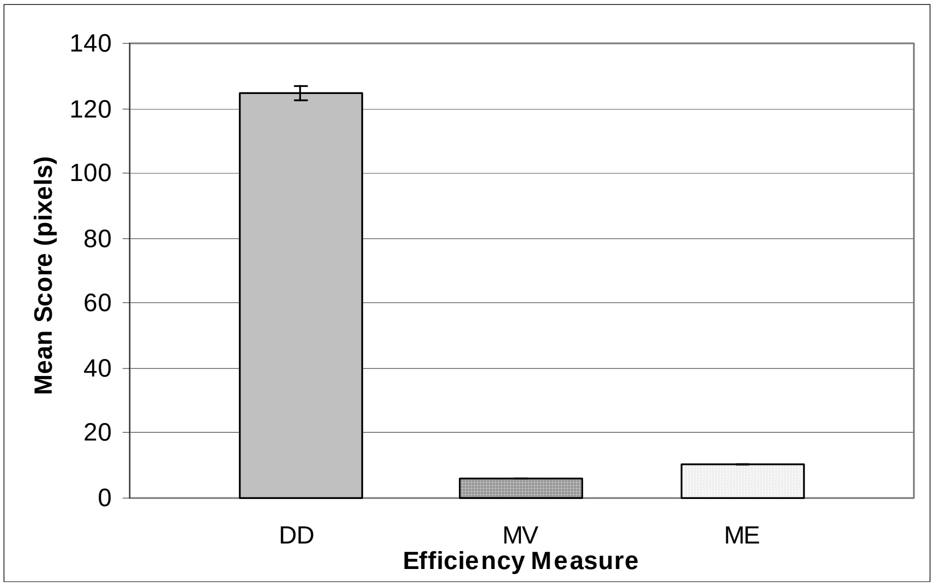


Figure 3.3. Summary of pixel-based efficiency measures.

Based on preliminary considerations of the data, the outcome measures captured in the investigation do not meet assumptions required to use regression analyses or other parametric statistics. That is, the distributions of the error terms for the measures were non-normally distributed and the sample sizes were not large enough to evoke the conventions of central limit theorem. In order to meet the assumptions of regression analysis, when necessary, transformations were made for each measure, and further outlying cases were identified and removed to strengthen the model (Neter, Kutner, Nachtsheim and Wasserman 1996).

Analyses: TT, VST, MT, DD, MV, ME

Forward stepwise linear regressions were used to analyze the contributions of the predictor variables to the overall variance on each efficiency metric, excluding FTHT and TTHT (which used logistic regression). A separate regression was generated for clinically-acquired factors and non-clinically-acquired factors for each measure. As stated, in order to meet the assumptions required by regression analysis, transformations were applied to each measure. These transformations and their interpretations are provided in Table 3.6.

The distributions for both FTHT and TTHT revealed that the majority of the participants, on the majority of the trials, generated target highlight times that were very small in magnitude. In fact, the 95th percentile for FTHT accounted for .0019% of TT, and for TTHT 1.1% of TT. The distributions for FTHT and TTHT did not meet the assumptions for a regression analyses, even with several transformation attempts. Instead, a logistic regression was applied the highlight measures. A more detailed description of the highlight measure analyses follows the linear regressions developed for the other efficiency measures.

Table 3.6. Description of efficiency measure transformation and an interpretation of the transformed variable.

Efficiency Measure	Transformation	Interpretation
TT	1/√TT	The rate of trial completion: As 1/√TT increases, the trial was completed at a faster rate.
VST	1/√VST	The rate visual search termination: As 1/√VST increases, the visual search completed at a faster rate.
MT	lnMT	The amount of time allocated to movement of the card to the drop pile: Higher lnMT indicates movement time to the icon was slower.
DD	lnDD	The distance traveled with the card on the way to the drop pile: An increased lnDD indicates movement time to the drop pile with the icon the icon covered a larger area, and was less efficient.
MV	√MV	The standard deviation in the distances of the path taken from the task axis from the mean on the movement of the card icon to the drop pile: An increase in √MV translates to a higher variability and less efficient movement.
ME	√ME	A measure of the average distances of the path taken from the task axis from the mean during the movement of the card icon to the drop pile: The higher the √ME the larger the distance and the less efficient the movement.

Results: Clinically-acquired Ocular Factors

Table 3.7 presents the model summary for six efficiency measures (excluding FTHT and TTHT), including R^2, R^{2-}adjusted, the standard error of R^2 and the Durbin-Watson statistic for each model. R^2 is an indication of the how much variance in the dependent model is explained by the predictor variables, their coefficients and the constant. R^2-adjusted corrects for a sometimes over-optimistic R^2 value. That is, it corrects the value to report an estimate more representative of the model for the true population. What constitutes a ‘good’ R^2 or R^2-adjusted value is subject to debate, and in fact, depends on the domain from which the model originates (Field 2000) Considering the variability inherent to human behavior, the R^2 and R^2-adjusted values for HCI investigations for older adults with visual impairments are somewhat liberally defined considering the specific domain and predictor variables (Pallant 2003). The Durbin-Watson statistic provides an indication of the correlation between errors for the predictor variable. This verifies that the model generated meets the regression assumption of independent residuals. Durbin Watson values close to 2 are desirable, and those values less than 1 or greater than 3 trigger concern for violations of the independent residual assumption (Field 2000).

With the exception of MV and ME, considering the high variability in human performance data, particularly for older adults, the emergent models were all good fits of the data, accounting for between 47 to 58% of the variability, based on R^2-adjusted. Durbin-Watson statistics for all six efficiency measures, were at acceptable levels (not <1 or > 3, and close to 2). For MV and ME, while the models demonstrated acceptable fits to the data, the predictors considered did not adequately account for the variability observed in either measure. Close to 90% of the variability in MV and ME remained unaccounted for in the models generated. It is therefore assumed that other factors, not integrated into this model, were largely driving these two measures of movement variability, and are not considered in the further exploration of the models. This is the first time MV and ME have been considered in the evaluation of interaction for the aging population (previous studies focused on input device with a normal population (e.g., MacKenzie, Kauppinen and Silfverberg 2001). The present investigation demonstrated little validation for the future incorporation of these measures, as the remaining terms were more informative of the relationships with the task and the predictor variables of interest in the hypotheses established.

Table 3.7. Clinically-acquired model summary for efficiency measures.

Statistic	1/√TT	1/√VST	lnMT	lnDD	√MV	√ME
N	2011	2011	1990	2004	2018	2028
R^2	0.58	0.518	0.487	0.473	0.345	0.326
R^2-adjusted	0.578	0.515	0.485	0.47	0.119	0.106
Durbin Watson	1.724	1.85	1.54	1.87	1.991	2.01

A linear model was generated for 1/√TT, 1/√VST, lnMT, and lnDD. From these models, enabling the value of the dependent measure to be calculated at various levels of the predictor variables using the following equations:

1/√TT = -.00115 -.00160 *SS* + .0000174 *Trial #* + .0000927 *Age* +.000103 *SF12MCS* -.0000373 *SF12PCS* + .000357 *Dexterity* - .00330 *NVA* + .000380 *CS* - .00279 *AMD Score* + .000270 *AMD Score*SS*

$1/\sqrt{VST}$ = .00669 - .00286 *SS* - .000813 *Column* +.000124 *Age* + .000188 *SF12MCS* - .000119 *SF12PCS* +.00314 *CS* -.00490 *AMD Score* + .000572 *AMD Score*SS*

lnMT = 10.260 + .115 *Row* - .002 *Trial#* - .008 *Age* - .010 *SF12MCS* + .887 *NVA* - .0806 *CS* + .166 *AMD Score* + .017 *AMD Score*SS* - .129 *AMD Score*AF*

lnDD = 3.81 +.244 *ISp* + .450 *Row* - .0265 *Drop Location* - .0308 *Dexterity* - .176 *NVA* + .0176 *CS* + .0783 *AMD Score* - .0590 *DD*

While the predictive models generated are of practical utility in estimating the actual efficiency measures, it is important to consider the predictive factors excluded from each model, in addition to the standardized Beta values. Table 3.8 provides an overview of the factors included in the model for each measure, the B, S.E. of B and B-std. The standardized coefficient (B-std) proves extremely beneficial to the comprehension of the models (Field 2000). It provides the means by which to quantitatively compare the relative impact of each predictor on the efficiency of each phase of the interaction.

Figure 3.4a-d provides summaries of the relative impact of each predictor variable. Variables with bars extending to the left of the 0 line imposed a decrease on the model of that measure, and to the right imposed an increase. Because this graph plots the standardized coefficient (B-std), relative comparisons can be made in terms of ‘how much more’ a predicator influences each model. When interpreting the direction of impact, an increase in $1/\sqrt{TT}$ and $1/\sqrt{VST}$ equates to a faster, improved time, while predicted increases in lnMT and lnDD represents slower movement times and indicates longer distances (i.e. diminished efficiency). Each figure enables the extrapolation on the cumulative affect of the predictors on each outcome measure, and can enable the consideration of different scenarios for various participant abilities and task factors.

Table 3.8. Summary of predictor variables for efficiency measure regression with clinically-acquired ocular factors (***** indicates that factor was excluded from the model).

TASK-RELATED FACTORS					
Variable		**1/√TT**	**1/√VST**	**lnMT**	**lnDD**
Constant	*B*	-0.00115	0.00669	10.26	3.81
	SE	0.00131	0.0023	0.147	0.0989
	p	0.38	0.004	<.001	<.001
SS	*B*	-0.0016	-0.00286	*****	*****
	SE	1.03E-04	1.87E-04		
	B-std	-0.319	-0.339		
ISp	*B*	*****	*****	*****	0.244
	SE				0.011
	B-std				0.374
AF	*B*	*****	*****	*****	*****
	SE				
	B-std				
Column	*B*	*****	-8.13E-04	*****	*****
	SE		9.5E-06		
	B-std		-0.132		
Row	*B*	*****	*****	0.115	0.45
	SE			0.012	0.013
	B-std			0.179	0.532
Drop Location	*B*	*****	*****	-0.057	-0.0265
	SE			0.007	0.0079
	B-std			-0.14	-0.0546
Trial #	*B*	1.74E-05	*****	-0.002	*****
	SE	1.28E-06		0.0002	
	B-std	0.134		-0.172	

Table 3.8. continued.

CLINICALLY-ACQUIRED VISUAL FACTORS					
Variable		**1/√TT**	**1/√VST**	**lnMT**	**lnDD**
NVA	*B*	-0.0033	*****	0.887	-0.176
	SE	3.33E-04		0.04	0.048
	B-std	-0.218		0.531	-0.088
CS	*B*	3.80E-04	3.14E-04	-0.08	0.018
	SE	2.7E-05	3.9E-05	0.003	0.004
	B-std	0.373	0.184	-0.72	0.131
AMD Score	*B*	-0.0028	-0.0049	0.166	0.078
	SE	1.58E-04	2.82E-04	0.017	0.011
	B-std	-0.723	-0.754	0.384	0.15
AMD * SS	*B*	2.70E-04	5.72E-04	0.017	*****
	SE	6.9E-05	1.25E-04	0.007	
	B-std	0.163	0.205	0.09	
AMD * ISp	*B*	*****	*****	*****	*****
	SE				
	B-std				
AMD * AF	*B*	*****	*****	-0.13	-0.059
	SE			0.01	0.012
	B-std			-0.23	-0.088

Table 3.8 continued.

PARTICIPANT RELATED FACTORS					
Variable		**1/√TT**	**1/√VST**	**lnMT**	**lnDD**
Age	*B*	9.27E-05	1.24E-04	-0.01	*****
	SE	8.60E-06	1.53E-05	0.001	
	B-std	0.213	0.17	-0.18	
SF-12 MCS	*B*	1.03E-04	1.88E-04	-0.01	*****
	SE	9.30E-06	1.67E-05	0.001	
	B-std	0.254	0.274	-0.24	
SF-12 PCS	*B*	-3.73E-05	-1.19E-04	*****	*****
	SE	8.07E-06	1.46E-05		
	B-std	-0.085	-0.162		
Dexterity	*B*	3.57E-04	5.56E-04	*****	-0.0308
	SE	3.54E-05	6.39E-05		0.0035
	B-std	0.262	0.243		-0.172

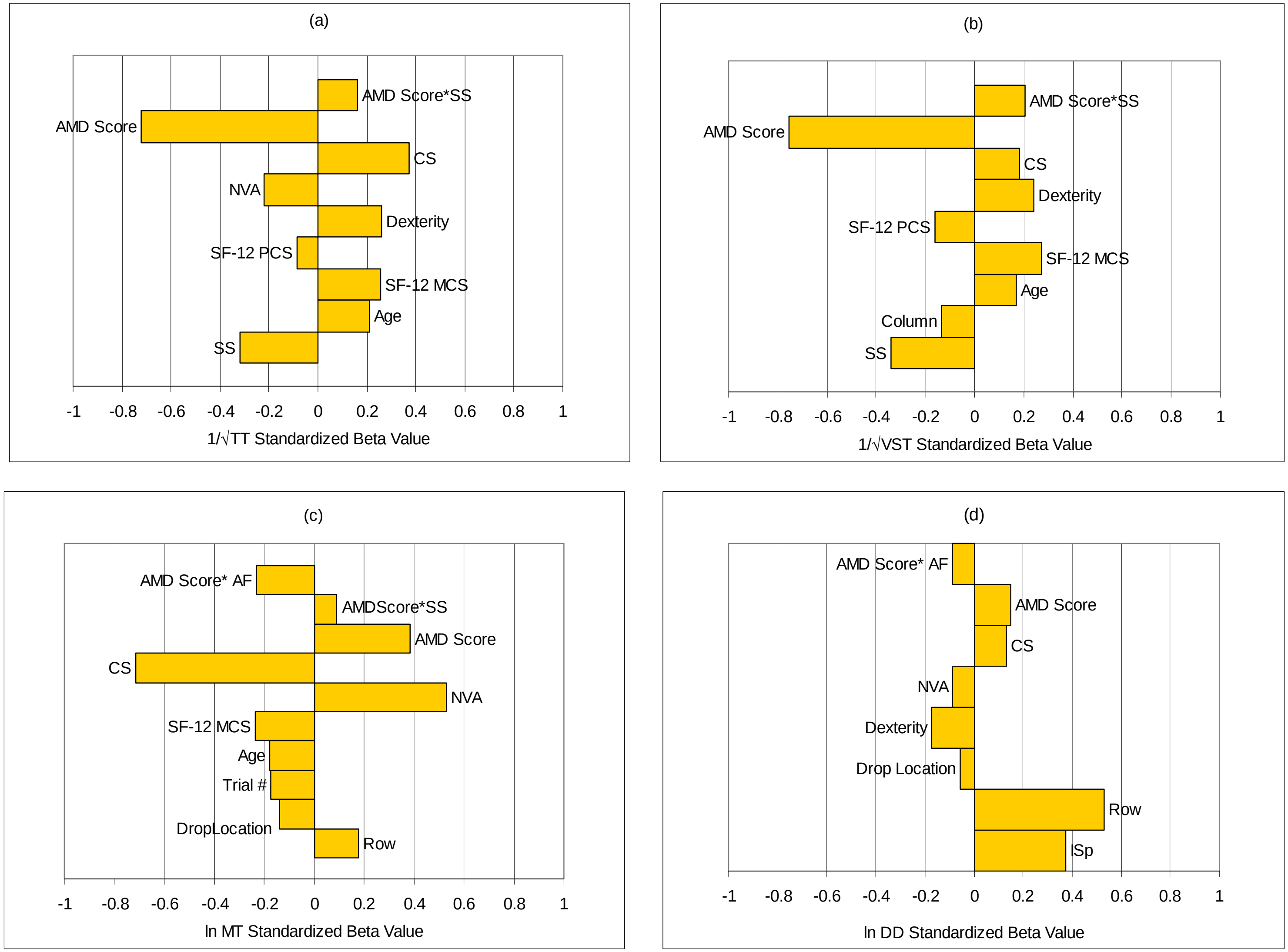


Figure 3.4 a-d. The relative impact of clinically-acquired predictor variables, illustrated via B-std for the accuracy measures.

Results: Non-Clinically-acquired Ocular Measures

The analyses of the efficiency measures were replicated using non-clinically-acquired ocular factors (i.e. VSAT percentile and overall VFQ score) in lieu of the clinically-acquired (CS, NVA, and AMD Score), maintaining the inclusion of the personal and task-related factors. Dependent measures were transformed for consistency with the previous clinical models. Linear regressions were generated for TT, VST, MT, and DD, while logistic regression was applied to the target highlight time metrics (TTHT and FTHT).

Table 3.9 presents the non-clinically-acquired ocular factor model summary for all efficiency measures (except MV and ME), reporting R^2, R^{2}-adjusted the standard error of R^2, and the Durbin Watson Statistic for each model. These models represent adequate to good fits of the data, and accounted for between 35.1% to 49.1% of the variability in the data set, with Durbin-Watson statistics below 3 and above 1.

Table 3.9. Non-clinically-acquired model summary for efficiency measures.

Statistic	**$1/\sqrt{TT}$**	**$1/\sqrt{VST}$**	**lnMT**	**lnDD**
n	2026	2022	1991	2009
R^2	0.493	0.434	0.354	0.453
R^2-adjusted	0.490	0.431	0.351	0.451
Durbin Watson	1.409	1.555	1.213	1.794

Based on the models, a linear equation was generated for each efficiency measure, the components summarized in Table 3.10. The predictive equations for modeling 1/√TT, 1/√VST, lnMT, and lnDD were as follows:

1/√TT = .0212 -.000881 *SS* + .00512 AF - .0002902 *Column* + .00023718 *Drop Location* + .0000117 *Trial#* -.0000794 *Age* -.000114 *SF12MCS* + .002816 *Dexterity* + .000115 *VSAT* +.00003092 *VFQ* - .0000534 *VFQ*AF* - .0000126 *VSAT * SS* - .00000916 *VSAT*ISp*

1/√VST = .0374 -.00141 *SS - .000818 Column* + .0000179 *Trial# -* .000145 *Age* - .000183 *SF12MCS* +.00307 *Dexterity* + .000210 *VSAT* - .0000168 *VSAT * ISp - .0000243 VSAT * AF*

lnMT = 7.680 -.700 *AF* + .1323 *Row* - .0578 *Drop Location* - .00136 *Trial #* + .00444 *Age* + .00713 *SF12MCS* - .00841 *SF12PCS* -.0324 *Dexterity* -.00339 *VSAT* +.00647 *VFQ * AF*

lnDD = 4.148 + .243 *ISp* + .0186 *Column* +.410 *Row* -.0283 *Drop Location* + .00361 *Age* - .00524 *VSAT*

Table 3.10. Summary of predictor variables for efficiency measure regression with non-clinically-acquired ocular factors (***** indicates that the exclusion of that predictor for that model).

TASK-RELATED FACTORS					
Variable		**1/√TT**	**1/√VST**	**lnMT**	**lnDD**
Constant	*B*	0.0212	0.0374	7.68	4.15
	SE	0.00110	0.00170	0.129	0.103
SS	*B*	-0.000881	-0.001408	*****	*****
	SE	0.000130	0.000230		
	B-std	-0.174	-0.166		
ISp	*B*	*****	*****	*****	0.243
	SE				0.0108
	B-std				0.371
AF	*B*	0.00512	*****	-0.700	*****
	SE	0.000895		0.106	
	B-std	0.618		-0.770	
Column	*B*	-2.90E-04	-0.000818	*****	0.0186
	SE	5.89E-05	1.04E-04		0.00793
	B-std	-0.0785	-0.132		0.0388
Row	*B*	*****	*****	0.132	0.410
	SE			0.0119	0.0128
	B-std			0.201	0.529
Drop Location	*B*	0.000237	*****	-0.0578	-0.0283
	SE	6.015E-05		0.00753	0.00807
	B-std	0.0629		-0.139	-0.0582
Trial #	*B*	1.167E-05	1.79E-05	-0.00136	*****
	SE	1.48E-06	2.50E-06	1.837E-04	
	B-std	0.132	0.121	-0.139	

Table 3.10. continued.

NON-CLINICAL VISUAL FACTORS					
Variable		**1/√TT**	**1/√VST**	**lnMT**	**lnDD**
VSAT	*B*	0.000115	0.000210	-0.00339	*****
	SE	7.400E-06	1.310E-05	0.00369	
	B-std	0.776	0.845	-0.388	
VFQ	*B*	3.092E-05	*****	*****	-0.00524
	SE	9.545E-06			0.000684
	B-std	0.103			-0.131
VFQ * SS	*B*	*****	*****	*****	*****
	SE				
	B-std				
VFQ * ISp	*B*	*****	*****	*****	*****
	SE				
	B-std				
VFQ * AF	*B*	-5.343E-05	*****	0.00647	*****
	SE	1.045E-05		0.00123	
	B-std	-0.562		0.619	
VSAT * SS	*B*	-1.257E-05	*****	*****	*****
	SE	2.893E-06			
	B-std	-0.203			
VSAT * ISp	*B*	-9.16E-06	-1.68E-05	*****	*****
	SE	1.797E-06	3.20E-06		
	B-std	-0.147	-0.161		
VSAT * AF	*B*	*****	-2.434E-05	*****	*****
	SE		5.100E-06		
	B-std		-0.2346176		

Table 3.10. continued.

PARTICIPANT-RELATED FACTORS					
Variable		**1/√TT**	**1/√VST**	**lnMT**	**lnDD**
Age	*B*	-7.94E-05	-0.000145	0.00444	0.00361
	SE	9.73E-06	1.58E-05	0.00117	0.000974
	B-std	-0.180	-0.198	0.0919	0.0632
SF-12 MCS	*B*	-1.14E-04	-1.83E-04	0.00713	*****
	SE	8.40E-06	1.37E-05	0.00104	
	B-std	-0.278	-0.266	0.158	
SF-12 PCS	*B*	*****	*****	-0.00841	*****
	SE			0.00117	
	B-std			-0.173	
Dexterity	*B*	0.00282	0.00307	-0.0324	*****
	SE	5.12E-05	5.28E-05	0.00516	
	B-std	0.204	0.133	-0.213	

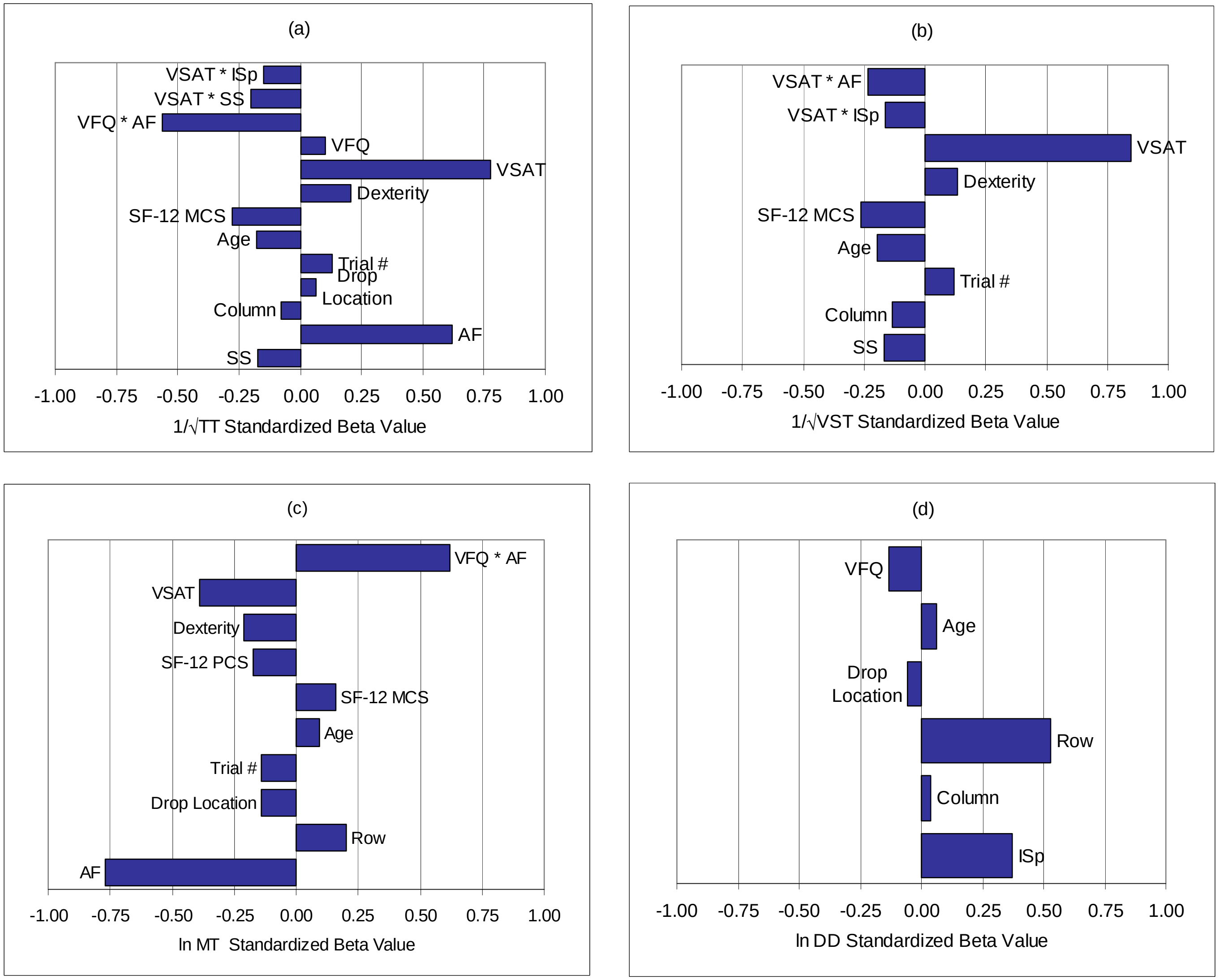


Figure 3.5a-d. The relative impact of non-clinically-acquired predictor variables, illustrated by B-std.

Figures 3.5a-d offer summaries of the relative impact of each predictor variable in these models using the non-clinically-acquired metrics. Variables with bars extending to the left of the 0 line imposed a decrease on the model of that measure, and to the right imposed an increase. Again, when interpreting the direction of impact, recall that an increase in $1/\sqrt{TT}$ and $1/\sqrt{VST}$ equates to a faster, improved time, while predicted increases in lnMT and lnDD indicates slower MT and longer distances respectively.

TT, VST, DD, MT Outcome Summary

Clinically-acquired Models

The outcomes of the regressions on TT, VST, DD and MT, revealed many interesting trends in the participants' interactions. They are discussed in the following section in terms of relevant outcomes, and will again be detailed at the conclusion of the chapter when the hypotheses are addressed.

Outcome #1: Independent Controlled Interface Variables

Spacing and Set Size. The interface-related independent variables influenced performance across all participants; the extent of influences dependent on the phase of the task. Logically, for the generated models, increases in SS triggered both slower VST and TT across all participants. Increases in ISp influenced longer DD across all the participants

<u>Outcome #2: Clinically-Acquired Visual Factors</u>

The persistent impact of clinically-acquired measures of ocular health and function on efficiency measures validates the dominant impact of the visual sensory channel on GUI-based tasks on platforms with small visual displays. Based on the standardized coefficients, *AMD score* and *CS* were reliable predictors for the models of all four measures (the only two predictors to be included in all four). This reaffirms the importance of investigations that focus on the impact of limited bandwidth of the visual sensory channel, for interaction on small visual displays.

AMD Severity Level. AMD Score was the most influential factor in both the TT and VST models. The models indicated that an increase in the severity of AMD can influence an increase in the VST and TT (slower rates of task completion and visual search termination). While not as influential on the icon dragging portion of the task, an increase in the severity of the AMD score imposed an increased movement time and longer dragging distance. The significance of this disease severity rating suggests there are implicit effects of AMD on functional vision, which are not effectively captured in the constructs of the functional vision metrics (i.e. NVA and CS) or the demographics.

This upholds previous work by Jacko in Colleagues (Jacko, Barnard et al. 2004; Jacko, Moloney, Kongnakorn, Barnard, Edwards, Emery et al. 2005)which concluded the same for participants with and without AMD working on a simple drag and drop task. This supports the continued use and exploration of ocular health assessments, an indirect measure of ocular function, as they can be highly indicative of the performance impediments experienced with GUI interactions.

Contrast Sensitivity. CS, like AMD Score, was included as a significant predictor for all four efficiency measures. CS was the second most influential factor on TT and the second most influential factor on DD. The model revealed improvements in CS to influence faster TT and VST. This is consistent with previous work which reported the significant impact of CS in the efficacy of interaction (Jacko 2000; Edwards, Barnard et al. 2005) and upholds the importance of this aspect of visual function in working with the computer.

Additional unanticipated results surfaced for CS in terms of the MT and DD. Improvements in CS influenced a faster MT but longer DD. Given that DD and MT are related to the same phase of the interaction, this result is interpreted as a *speed-accuracy* trade off. That is, while participants with better contrast sensitivity were fast with their use of the stylus to move the card to the drop pile, they were less accurate with respect to the efficiency of the path taken to the drop pile. Speed accuracy trade-offs are a common occurrence in human integrated systems, and have been observed in several domains that operate on discrete and continuous motor control, and specifically in the aging population (Fitts 1954; Pew 1969; Darling, Cooke and Brown 1989).

Near Visual Acuity. The final clinical measure of visual function, NVA, was included as a predictor in the models of TT, MT, DD, but not VST. As NVA worsened (the value approached 1) in the model, TT and MT were slower, but DD was prone to be shorter. A speed-accuracy trade-off was therefore observed for NVA on the measures of DD and MT. As NVA worsened, MT increased to indicate slower icon movement and the distance traveled was more likely to cover a shorter area.

This trade-off was consistent with that observed in CS: As vision degraded (in terms of CS and NVA) the interaction was affected by increased TT, VST (for CS only) and MT, but the path taken to the drop pile was shorter (DD). This suggests that the participant who experienced lower CS and NV had the aptitude to complete the task effectively (as seen with the decreased DD), but it took them longer to interact with the interface, and this also demonstrates that the individuals with the better vision, while they are prone to longer DD, this does not affect their MT or DD, nor does a DD alone indicate error-filled task behavior. One of their strategies for coping with their poor vision is likely to make sure their movement covers as small of an area as possible.

Additionally, for the visually health participants, while they are prone to demonstrate longer DD, the increase in distance traveled was not observed to detract from these participants' task accuracy or TT. Over the small area of the display, for participants without ocular deficiencies, the longer distance does not equate into global task efficiency decrements.

Outcome # 3: Interactions of Disease Severity and Interface Variables

*AMD Score * SS*. As AMD * SS increased (i.e. either AMD score increased and/or SS increased), trials were completed faster and visual search was terminated more quickly. This result at first consideration is suspect. It appears that the impact of spacing gives those individuals with more severe AMD an advantage. However, it is not without merit in consideration of the fact that the AMD Score construct is liable to encompass an indication of visual field interruptions (scotomas, aberrations, and distortions). The higher the severity rating of AMD, the more drusen are present on the macula. In addition, the higher severity scores are more indicative of the ‘wet’ form of AMD. Both the presence of drusen and wet AMD are indicative of escalating disruptions to the visual field. These interruptions in the visual field may in fact limit the number of icons that are viewable at once in the visual field. The inclusion of more icons might not affect these participants or the method by which they scan the interface because they do not perceive much difference at the display onset from smaller SS conditions.

*AMD*AF.* The impact of AF on the task did not have a measurable impact across all of the participants in these models. However, the interaction between AMD*AF proved very interesting. An increase in the AMD*AF interaction term (i.e. AF was present, and AMD was present and or increased in severity), prompted a faster MT shorter DD, the phases of the task to which the feedback directly applies. This is a compelling result, as it demonstrates that the inclusion of supplemental non-visual cues can counteract the negative effects imposed on the interaction by the disease and is more influential at intercepting the efficiency issues at the more severe levels of AMD. These gains from the inclusion of AF were not observed for those without ocular diagnosis (AMD Score = 0).

This gain is quantifiable, through the examination of the standardized coefficient values. For MT, the B-std for AMD Score was -.384 and for AMD*AF, B-std = .15. That said, according to the model, if a person with an AMD Score of 1 would experience a 39% improvement in MT with the inclusion of auditory feedback (all other factors held constant). In terms of DD, B-std for AMD Score = -.229, and AMD * AF = .088. In this model, a person with AMD Score of 1 would experience a

38% reduction in the distance travelled with the card on the way to the drop pile, with the inclusion of AF (all other factors held constant)- all notable performance gains.

These results are especially notable because they clarify previous, contradictory results on the utility of AF for a drag and drop. Recalling that in their investigation of multimodal feedback for a single file to single folder drag and drop (simple) demonstrated that those conditions using auditory feedback benefited the participants with AMD of a range of visual acuities, particularly those within stratified groups with the most severe visual dysfunction (e.g., Jacko, Emery, Edwards et al. 2004). In contrast, Vitense, Jacko and Emery observed for a different auditory sound, in a complex drag and drop task (multiple files and folders), to actually inhibit performance times (Vitense, Jacko et al. 2002, 2003). The evidence presented by the results of the thesis suggests that AF had a positive influence on performance, but only in measurable amounts as the severity of the disease worsened.

Outcome # 4: Assorted Interface and Task Aspects

While additional display and task features impacted the efficiency performance models, their influence was substantially less in magnitude compared to the visual factors. The models accounted for the variability in Drop Pile location, target card icon Column and Row, and the Trial # in the regression. While these factors are important in the generation of robust models, they were of less interest in the context of the overall dissertation as they were not deliberately controlled independent variables. Furthermore, the effect of each was not surprising, and for the majority of these attributes, the impact was slight, as compared to the impact the other predictors.

Trial #. The consideration of Trial # in the models served to expose any learning effect for participants' interactions from trial 1 to trial 162. The learning effect was included as predictors of performance in the model of MT only. MT decreased as the trial number increased. This supports the fact that the participants gained skill in their use of the stylus and handheld display over the trials, while they were consistently challenged in the visual search portion of the task between trials.

Target Icon Card Column Location. Logically, target card icons located further to the right of the grid imposed longer VST. This is consistent with the nature of visual scan for Western users, who work from left to right to locate an icon, and also reflects that these columns were furthest from the column of drop piles. Those icons located in rows lower on the display imposed longer MT and DD.

Target Card Icon Row Location. Surprisingly, Row was one of the most influential predictors of DD. Those icons lower on the display resulted in much longer DD.

Target Drop Pile Location. Drop location was accounted for in the models of MT and DD. The model showed a potential for shorter DD and MT as the drop pile was relocated at lower positions on the display.

These results suggest that the location of the targets can influence the ease of visual tracking and stylus use, and warrant further exploration with this population, input devices, and screen layout.

<u>Outcome #5: Personal Factors</u>

Several personal factors were included as predictors in the efficiency models, but did not dominate any of the models.

Age. The effects of age were unexpected in these models, as increases in age imposed faster TT, VST and MT. Two possible explanations for this include the fact that clinical vision metrics account for several aspects of age in the model. Additional explanations for this effect can be attributed to the nature of the task employed in this study. It could be the case that the older participants had more recent experience playing cards, and were able to locate the target icons more quickly than those who play cards less frequently. Older participants may have more experience playing cards than the younger participants, and likely had more spare time for such activities (the majority of "young-old", e.g., ~50-65 yrs participants were not yet retired).

The use of familiar icons can increase users' comfort level, and thus proficiency, with new technologies. This is in opposition to the outcomes of the regressions generated by Edwards and colleagues. However, the population considered in that work was diagnosed with Diabetic Retinopathy, a disease affecting a wider range in age (Edwards, Barnard et al. 2004; Edwards, Barnard et al. 2005). That said, the current results provide explicit insight into the older adult population, and how "young-old" (50-65) individuals differ from those considered part of the "older" (>65 yrs segment.

Dexterity. As the manipulation of icons on the handheld display using the stylus is largely a visual-motor coordination task, it is not surprising that improvements in Dexterity influenced the faster TT and shorter DD. Dexterity's influence on VST was unexpected. Improvements in dexterity, in the model, led to faster VST. This may be an indication that the participants were using the stylus to guide their search for the icon, utilizing it as a pointer, or placeholder.

Mental and Physical Health. The final personal attributes, SF-12 MCS and PCS were also included in the models for task efficiency. Increases to MCS (the mental health component) generated an increased or faster rate for TT, VST and MT. PCS, however, emerged in less rational patterns. Increases in PCS in the model generated slower TT and VST. While this is contradictory to what is expected, it is important to remember that this is the physical health as rated by the participants in an interviewer-administered survey.

Non-Clinically-acquired Models

Outcome#1: Differences between Clinical and Non-Clinical Models

Not unexpectedly, the models for the non-clinically-acquired visual factors differed from those generated for the clinically-acquired factors. This is a result of the discrepancies in how the variability in each measure was accounted for by the non-clinical measures as opposed to the clinical measures. Even so, several of the emergent trends in the predictors included in both were consistent between the models with a handful of new clarifying trends. Also, the personal factors of age, MCS measures demonstrated inverse effects in these models from their behavior in the clinical models. Table 3.11 summarizes how the models generated under the clinical and non-clinical predictors differed. The *Consistent* column identifies those predictors which where included in the equations under both conditions, with the same impact on the outcome variable. The column labeled *Unique to Non-Clinical Models* are predictors that were included in the clinically-acquired model for that measure, but not the non-clinically-acquired, and likewise for the column labeled Unique to Non-Clinical Models. Finally, the column that designates *Reverse Effect* designates predictors that appeared in the equations for both sets, but the impact on the dependent variables was in the opposite direction.

While the inclusion of additional terms is not a concern, predictors having an inverse influence on the outcomes between the clinical and non-clinical models merit deliberate consideration (but are not entirely surprising). These 'Reversed Effects' were limited to the personal factors of Age and SF-12 MCS and they accounted for some of the variability in the dependent variables, but limited influence relative to the other included predictors. A plausible explanation for this reversal is that in the different constructs actually included in the different measures. That is, the clinically-acquired variables may account for the most relevant aspect of age and health that are actually not accounted for through the VSAT or the VFQ. In this study, regression simultaneously including the VSAT, VFQ and the clinical factors was not reported, due to the violation on the multi-colinearity rule for linear regression. Future work should take into account the consolidation of both classes of visual assessment.

Table 3.11. Deviation of predictor variables between clinical and non-clinical models.

Model	Consistent	Unique to Clinical Models	Unique to Non-Clinical Models	Reverse Effect
1/√TT	SS Dexterity	AMD Score CS NVA AMD Score*SS	VSAT VFQ VSAT*ISp VSAT*SS VFQ*AF AF Trial # Drop Location Column	SF-12 MCS Age
1/√VST	SS Dexterity Column	AMD Score CS NVA AMD Score*SS SF-12 PCS	VSAT VSAT*ISp VFQ*AF Trial #	SF-12 MCS Age
lnMT	Trial # Drop Location Row	AMD Score CS NVA AMD Score*SS AMD Score*AF	VSAT VFQ*AF AF Dexterity SF-12 PCS	SF-12 MCS Age
lnDD	ISp Drop Location	AMD Score CS NVA AMD Score*AF Dexterity	VFQ Column Age	None

<u>Outcome #2: Independent Controlled Interface Variables</u>

Set Size and Icon Spacing. Increases to SS were shown to influence slower VST and TT, consistent with the clinically-acquired models. Also consistent was the impact of larger ISp on increased DD. The relative impacts of these terms on the models (B-std) were also equivalent in comparison to the other predictors included in each model.

Auditory Feedback. Unlike the models derived from the clinical factors, AF emerged as a significant influence on two models – TT, and MT. The presence of AF was the most heavily weighted factor in the model of MT and the 2nd most heavily weighted factor in the model of TT. AF had a sizeable positive influence on the models; the presence of AF generated decreased TT and MT. However, caution must be taken in this interpretation - while it appears that this implies that AF had positive influence across all participants, the interaction term of VFQ*AF and VSAT*AF tell a different story. In other words times increase with the presence of AF and/or improved visual attention or perceived visual functioning scores, effectively canceling the positive role of the feedback. Only those individuals with lower scores on these non-clinical tests can, according to the model, experience the benefits of AF.

<u>*Outcome #3 Non-Clinically-acquired Visual Factors*</u>

Visual Attention. The VSAT percentile score was weighted significantly in the models of TT, VST, MT, but not on DD. In the prior three, improvements in visual attention were influential on increased rates of TT, VST and MT. In fact, VSAT percentile was the most heavily weighted B-Std in the models of TT and VST. This confirms that visual attention; a combination of visual function and adaptation to disease is a critical determinant of efficacy in a visually intensive interaction task, and a global measure of task efficiency.

Perceived Visual Functioning. Increases in the VFQ-Overall score (i.e. improved perception of visual function on daily activities) in the model generated faster TT and DD. Notably, the relative weight of VFQ Overall was significantly less than the VSAT in its impact of TT (VFQ B-std = .103, VSAT = .776). However, VFQ Overall was also included in model for DD, in which VSAT was not included. In the model of DD, VFQ had a median B-std value compared to the other terms. VFQ was also included as an interaction with AF in the model of MT and weighted almost twice that of the VSAT B-std. This could be an indication that the VSAT may be more appropriate for detecting global measures of task performance, while the VFQ may more effectively capture the ability of the individual's capacity for coordination tasks, such as moving a stylus and visually tracking movement simultaneously.

Outcome #4: Interactions between VSAT, VFQ and Interface Variables

*VSAT*ISp.* Unique to the models with the non-clinical vision factors, the inter-icon spacing and VSAT score interaction significantly influenced TT and VST. The VSAT*ISP factor increased in the model, the rate of TT and VST increased, or was slower. This can be interpreted that the negative effects of increased ISp were amplified for those with better levels of VSAT (if both VSAT and ISp increase).

The impact of ISp is not observed across all participants. It may also be the case that increasing spacing, for those individuals with lower VSAT scores, may actually negate the impacts of poor visual attention. This is logical, as the increased proximity of icons can lead to an individual being more distracted in his or her search by the immediacy of adjacent icons in the visual field. However the magnitude of impact of the VSAT*ISp term in both TT and VST is relatively small (TT B-std = -.147, VST B-std= -.161) compared to the impact of VSAT in TT and VST (TT B-std = .776, VST B-std = .845).

*VSAT * SS.* The interaction between the number of icons on the display (SS) and the VSAT were included in the prediction of TT. Similar in terms of direction and the magnitude of the effect of VSAT*ISP, increases in the value of the interaction term yielded a slower TT. That said, the performance gap generated by the VSAT is effectively lessened through increasing the number of icons on the display. Those individuals with the best levels of visual attention, according to the model, experienced a decrease in the rate of TT greater than that experienced by individuals with lower or worse VSAT scores.

*VSAT * AF.* The interaction between AF and VSAT was included as a predictor of VST. The inclusion AF influenced a slower VST with increasing, or improved VSAT scores. AF did not emerge as providing a significant main effect in this model, suggesting that across all participants, there was a slower VST in the presence of AF. The degree of influence on VST was proportional to the improvements in visual attention. Participants with the best visual attention scores were more easily distracted by the feedback, even during phases of the task that were unrelated to AF.

*VFQ*AF* was significant in the model of MT (like the clinical models), but unlike the clinical models, it was significant in its influence on TT as well. As stated before, the inclusion of supplemental non-visual feedback, while it showed signs of improving TT and MT for all participants, the interaction between VFQ*AF effectively dampened this performance gain when AF was present, and VFQ increased. This is realized through the comparison of the B-std for the models of TT and MT. For TT B-std AF= .618, VFQ*AF = .562; for MT, B-std for AF= .618; for VFQ*AF= .619. The values quantifying the impact of each term are very close in proximity and effectively convey how powerful the non-visual cues can be to the enhancement of performance for those individuals with lower perceived visual function in terms of daily activities.

Analyses: TTHT FTHT

As stated, the distribution of the participant performance on both FTHT and TTHT were not amenable to a regression analyses, despite several transformation attempts. Instead, logistic regression was used to identify which factors most impact the probability of each highlight measure exceeding threshold values. These threshold values were derived from the distribution of the participant highlight scores, and designating a cut point at the 85^{th} percentile. This cut point was chosen because it created a valid distribution to which logistic regression was applied. In addition the results of the logistic regression can be interpreted in a meaningful way.

The outcome of the logistic regression can designate the predictors' influence on the probability that a target highlight time be classified above the 85^{th} percentile. The 85^{th} percentile for TTHT was calculated to be 16 msec, based on average participant performance, constituted .5% of a typical trial time. Likewise, the 85^{th} percentile for FTHT was identified as 3 msec, or .14% of the average trial time.

Forward-stepwise logistic regressions, based on the likelihood ratio statistic, were applied to both FTHT and TTHT for the models employing the clinically-acquired measures. Stepwise regression methods, have, in other statistical analyses and discussion, been identified as optimal for exploratory studies, in which little to no previous research exists in the area (Fahrmeir and Tutz 1994; Field 2000).These logistic regression models included predictor variables consistent with those linear regression models executed on the other efficiency measures. The logistic regression produced valid models for predicting the likelihood that either FTHT or TTHT would be exceeding the 85^{th} percentile for performance (i.e., the longest highlight times).

Results: Clinically-Acquired Visual Factors

The Hosmer and Lemeshow Goodness of Fit assessment (HL test) assessed the null hypotheses that there are no significant differences between the observed and predicted values of the dependent variable (e.g., it is desirable to have a p value greater than .05 to imply a good fit). The HL test was calculated not to significant for the model of TTHT (p = .962), but was significant for the model of FTHT (p =.025). This indicates the model generated for TTHT was a good fit, but that the model generated for FTHT was not. Furthermore, for this sample population the TTHT model was correct in its categorization on 85.6% of the cases and the FTHT model was accurate in 91.2% of its classifications. For FTHT, the discrepancy between the HL test and the accurate classification may be a result of sample size.

The coefficients, test statistics and significance levels are described in Table 3.12. Figure 3.6 illustrates the magnitude of the impact of each predictor in each model via Exp (B). Exp (B) is a very useful in the interpretation of the model because it is an indication of the ‘change in odds’ for the outcome measure, based on the given predictor increasing by one unit. Exp (B) values of less than one influence a decrease in the probability of the dependent variable and Exp (B) values greater or equal to one increase the likelihood of the outcome. The further a bar extends, in either direction, away from 1, the greater the impact on changing the probability for errors to occur.

Table 3.12. Significant variables associated with logistic regressions on FTHT and TTHT using clinically-acquired measures (***** indicates exclusion of that predictor from the model).

TTHT				
Variables	**Coefficients (B)**	**Wald Statistic**	**p**	**Exp (B)**
Constant	-0.50	1.51	0.22	0.61
SS	*****			
ISp	-0.57	45.33	<.001	0.568
AF	0.39	4.04	0.04	1.483
Column	*****			
Row	-0.485	20.39	<.001	0.615
Drop Location	-0.356	32.99	<.001	0.700
Trial #	-0.005	10.98	0.001	0.995
Age	*****			
SF-12 MCS	*****			
SF-12 PCS	*****			
Dexterity	-0.055	6.09	0.014	0.947
NVA	1.18	13.34	0.000	3.247
CS	*****			
AMD Score	0.55	43.06	<.001	1.729
AMD Score*SS	*****			
AMD Score* ISp	*****			
AMD Score* AF	-0.513	15.40	<.001	0.599

Table 3.12. continued.

FTHT				
Variables	**Coefficients (B)**	**Wald Statistic**	**p**	**Exp(B)**
Constant	-220.00	1.57	0.692	0.80
SS	*****			
ISp	*****			
AF	*****			
Column	*****			
Row	*****			
Drop Location	*****			
Trial #	*****			
Age	*****			
SF-12 MCS	-0.02	6.99	0.008	0.977
SF-12 PCS	*****			
Dexterity	*****			
NVA	-1.47	21.13	<.001	0.231
CS	*****			
AMD Score	*****			
AMD Score*SS	*****			
AMD Score* ISp	*****			
AMD Score* AF	*****			

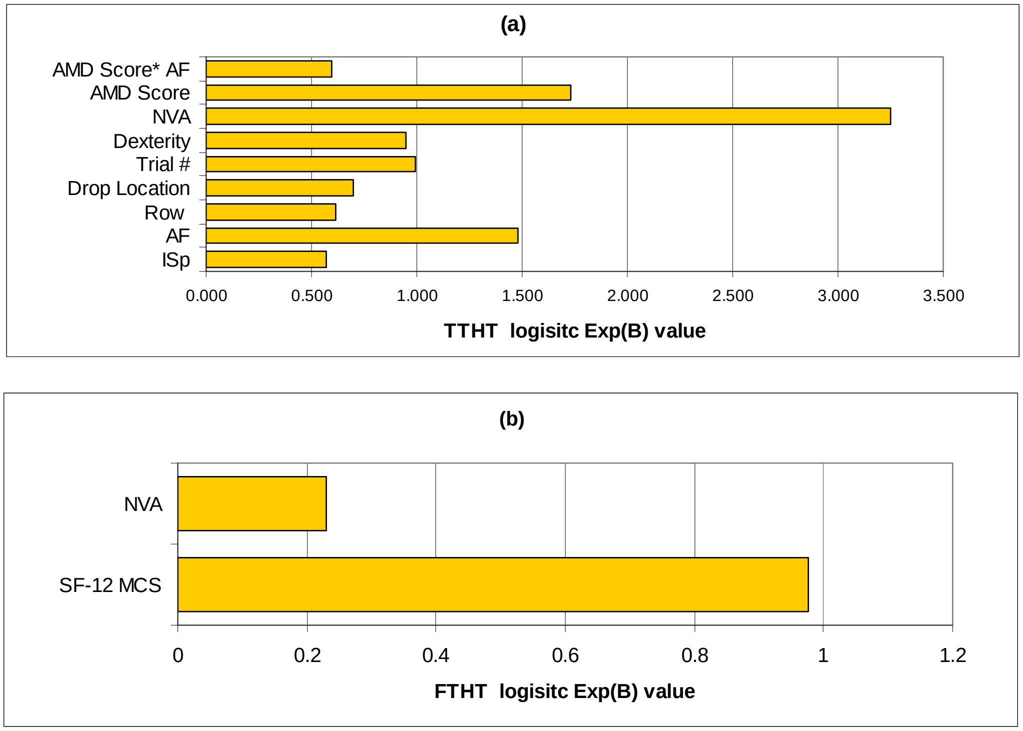


Figure 3.6. Illustration of the relative impact of TTHT and FTHT on the probability of highlight times excessive of the 85th percentile with clinically-acquired predictors.

Results: Non-Clinically-Acquired Visual Factors

Forward-stepwise logistic regressions, based on the likelihood ratio statistic, were applied to both FTHT and TTHT using the non-clinically-acquired factors, personal and task-related factors, consistent with the other models. The HL for TTHT logistic regression was not significant (p= .958) and correctly predicted the outcome in the sample population 85.4% of the time indicating a good fit and model. For FTHT, HL was significant (p<.001), but similar to the clinically-acquired factor model, a very high percentage of the cases in the sample population were correctly classified (91.2%). Table 3.13 describes the coefficients, test statistics and significance. Additionally, Figure 3.7 illustrates the magnitude of the impact of each predictor in each model with Exp (B).

Table 3.13. Significant variables associated with logistic regressions on FTHT and TTHT using non-clinically-acquired measures (***** indicates models predictors exclusion from a model).

TTHT				
Variables	**Coefficients (B)**	**Wald Statistic**	**p**	**Exp (B)**
Constant	0.19	0.87	0.351	1.21
SS	******			
ISp	******			
AF	******			
Column	******			
Row	-0.466	19.32	<.001	0.627
Drop Location	-0.33	29.444	<.001	0.719
Trial #	-0.004	8.16	0.004	0.996
Age	******			
SF-12 MCS	******			
SF-12 PCS	******			
Dexterity	******			
VFQ	******			
VSAT	******			
VFQ * SS	******			
VFQ * ISp	-0.007	48.29	<.001	0.993
VFQ * AF	******			
VSAT * SS	******			
VSAT * ISp	******			
VSAT * AF	******			

Table 3.13. continued.

FTHT				
Variables	**Coefficients (B)**	**Wald Statistic**	**p**	**Exp (B)**
Constant	0.39	0.35	0.56	1.477
SS	*****			
ISp	*****			
AF	*****			
Column	*****			
Row	*****			
Drop Location	*****			
Trial #	*****			
Age	*****			
SF-12 MCS	-0.026	6.24	0.01	0.974
SF-12 PCS	*****			
Dexterity	0.195	12.723	<.001	1.215
VFQ	-0.046	19.96	<.001	0.955
VSAT	*****			
VFQ * SS	*****			
VFQ * ISp	*****			
VFQ * AF	*****			
VSAT * SS	*****			
VSAT * ISp	*****			
VSAT * AF	*****			

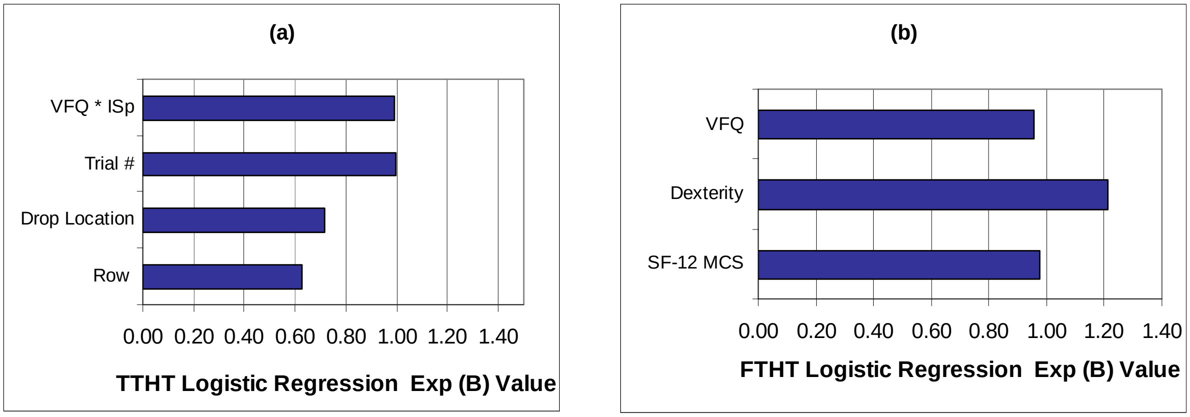


Figure 3.7. Illustration of the relative impact of TTHT and FTHT on the probability of highlight times excessive of the 85^{th} percentile with non-clinically-acquired predictors.

TTHT FTHT Outcome Summary

As the models generated for FTHT under both the clinical and non-clinical model were not good fits, the resultant logistic regression models and relationships are deemed not robust. Therefore, the models of TTHT were relied on to characterize the efficacy of the feedback and the participant's success in the 'drop' portion of the interaction.

Clinically-Acquired Visual Factors

Outcome # 1: Clinically-acquired visual factors

CS. In the logistic model of TTHT, AMD Score, NVA and the interactions were included as predictors, while CS was not. This was surprising, because the models of the efficiency metrics included CS on all four aspects of the task, with a particularly heavily weighted influence on MT. It is therefore deduced that contrast sensitivity was influential on the search and visual tracking of the icon during movement, but does not weigh in on the alignment/release of the icon into the drop pile.

AMD Severity Level and NVA. Increases to AMD Score, or as the severity rating of disease worsened, imposed an increased on the likelihood of the participants' TTHT to exceed the 85^{th} percentile. As NVA worsened from .1 to 1 it also increased the likelihood of TTHT exceeding the 85^{th} percentile in the logistic regression model. The impact of AMD Score was the second most influential factor on increasing the probability of TTHT extending beyond the identified threshold, while NVA was the most highly influential on the outcome. This could be an indication that the ability to resolve fine details, such as the point at which the card icon is accurately positioned over the drop pile, drives the efficacy of the task. Furthermore, this upholds previous work on the drag and drop and the population which stratified participants with AMD based on their visual acuity and demonstrated the longest highlight times for the worst visual acuity (Jacko, Scott et al. 2003; Jacko, Moloney, Kongnakorn, Barnard, Edwards, Emery et al. 2005).

Outcome #2: Independent Controlled Interface Variables

Auditory Feedback and Inter-icon Spacing. In the model, AF was shown to increase the likelihood of TTHT being in excess of the 85^{th} percentile across all participants. However, there was also a significant interaction between AMD Score and AF, which effectively decreased the likelihood of TTHT exceeding the 85^{th} percentile. Again this is an indication that AF can neutralize the effects of the disease, to make lessen the performance gap between the ranges of disease and the controls. However it is also an indication that AF can potentially inhibit the performance of those without AMD.

Increases in ISp in the model imposed a decrease in the probability of TTHT exceeding the 85^{th} percentile. This implies that the positioning of the card icon was easier in conditions with more space between the icons, even though increases in ISp increased DD. This suggests that larger spacing of the target destinations (e.g. folders) is more critical in the release portion of the task, while the decreased spacing of the target icons and distracter icons can optimize the initial movement of the icon as it approaches the destination, for all participants

Outcome #3: Manual Dexterity

Increases in manual dexterity were shown to decrease the likelihood of TTHT exceeding the 85^{th} percentile. Interestingly, dexterity in the other efficiency measures was included in the models of TT, VST, and DD but not MT. The influence of dexterity, based on the inclusion in the model of TTHT in addition to the other efficiency measures supports the fact that HCI is a visually rigorous task, which requires notable amounts coordination between the visual and motor systems.

Outcome #4: Other interface features

Consistent with the models on the other efficiency measures, there was a small learning effect for improved TTHT – trials later in the task influenced a lower probability of TTHT exceeding the 85^{th} percentile. Finally, the row of the target icon and target drop pile also imposed a decrease on the outcome probability. Again, this demonstrated that the participants experienced ease in the manipulation of the icons lower on the display, and merits further study of the impact of layout of items on a small display.

Non-Clinically-acquired Visual Factors

The logistic regressions based on the non-clinically-acquired visual factors did not illuminate the interaction as well as the clinical factors. The model included just four predictors which influenced a decrease in the likelihood of TTHT exceeding the 85^{th} percentile: 1) VFQ * ISP; 2) Trial #, 3) target drop pile location, and 4) target card Icon row.

The effects of Trial #, Drop Location, and Row were all consistent with the results of the clinical model, while the VFQ*ISp was unique to this model. The influence of this interaction term however was relatively small in terms of how much the probability decreased as the VFQ score and/or the ISp increased. However, it is consistent with the results on the other efficiency measures, for which there was an interaction between VSAT and ISP on both TT and VST. The non-clinically-acquired visual factors, for TTHT proved much less diagnostic than the clinically-acquired factors.

3.3 Accuracy

General Summary

The mean scores for each accuracy measure, as summarized in Table 14, reiterated the high level of success participants experienced in their task interactions on the handheld. The low error rate in this study is not surprising, as other investigations with similar populations (and no time limit on the task) showed similar indications of task efficacy (Jacko, Moloney, Kongnakorn, Barnard, Edwards, Emery et al. 2005). Figure 3.8 provides additional insight into those accuracy measures summarized in Table 3.14, highlighting the frequency with which each error occurred across participants. The errors principally occurred with a frequency of 1 or not at all. Accordingly, logistic regression was applied to the accuracy measures to determine which predictor variables increased the probability of one or more errors occurring in a single trial (Neter, Kutner et al. 1996).

Table 3.14. Summary of accuracy measures of handheld interaction, (abbreviations for each measure appear in italics).

Accuracy Measures	
Number of missed opportunities: Over no drop-*OND*	Mean (Std. Error): .31 (.015) Median: <.001 Minimum:<.001 Maximum: 8.00
Number of accidental drops-*AD*	Mean (Std. Error): .14 (.014) Median: <.001 Minimum: <.001 Maximum: 12.00
Number of task axis crossings: Icon dragging-*TX*	Mean (Std. Error): .56 (.020) Median: <.001 Minimum: <.001 Maximum: 10.00
Number of movement direction changes- *DC*	Mean (Std. Error): 1.13 (.046) Median: 1.00 Minimum: <.001 Maximum: 36.00

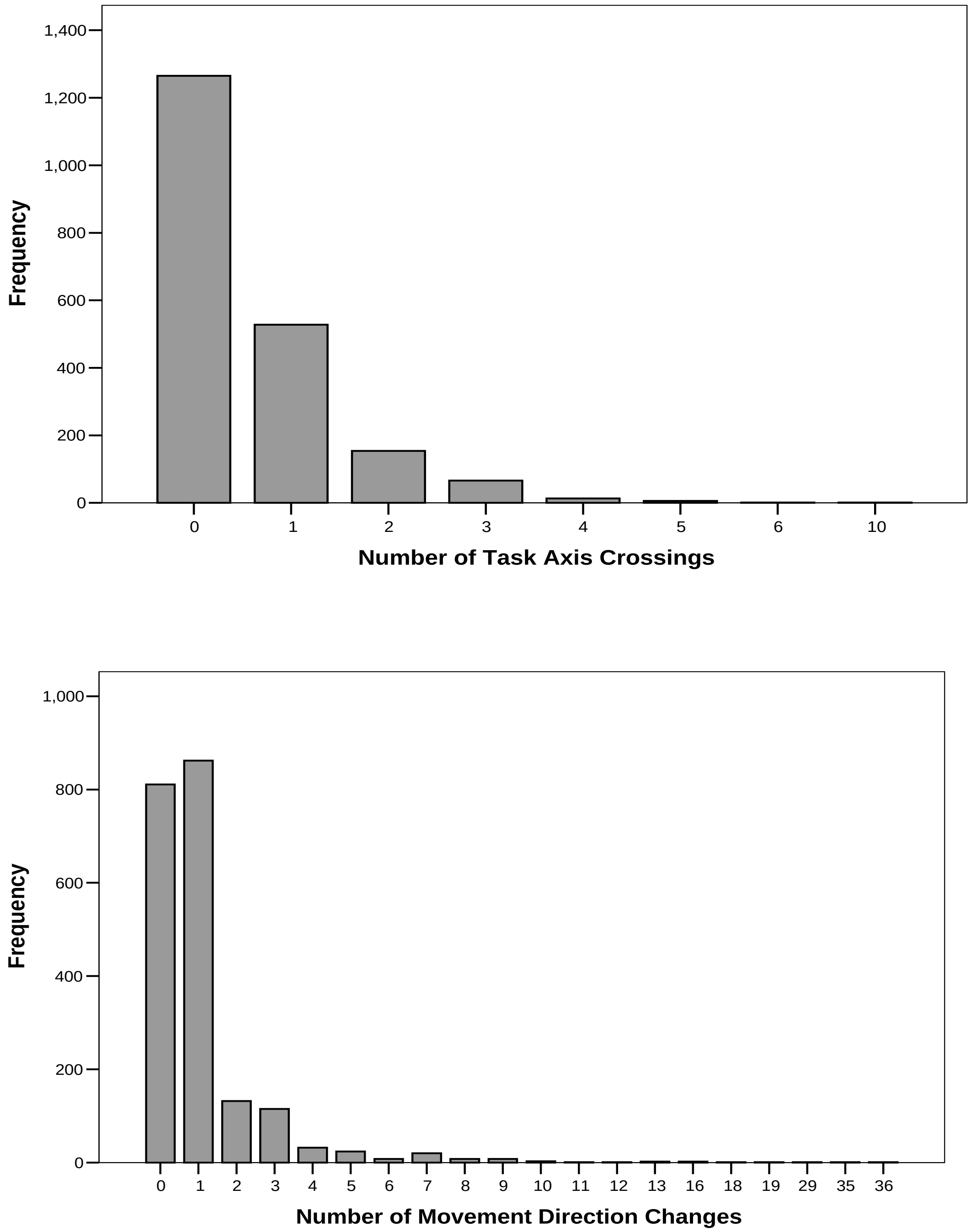


Figure 3.8. Frequency distribution of accuracy measures.

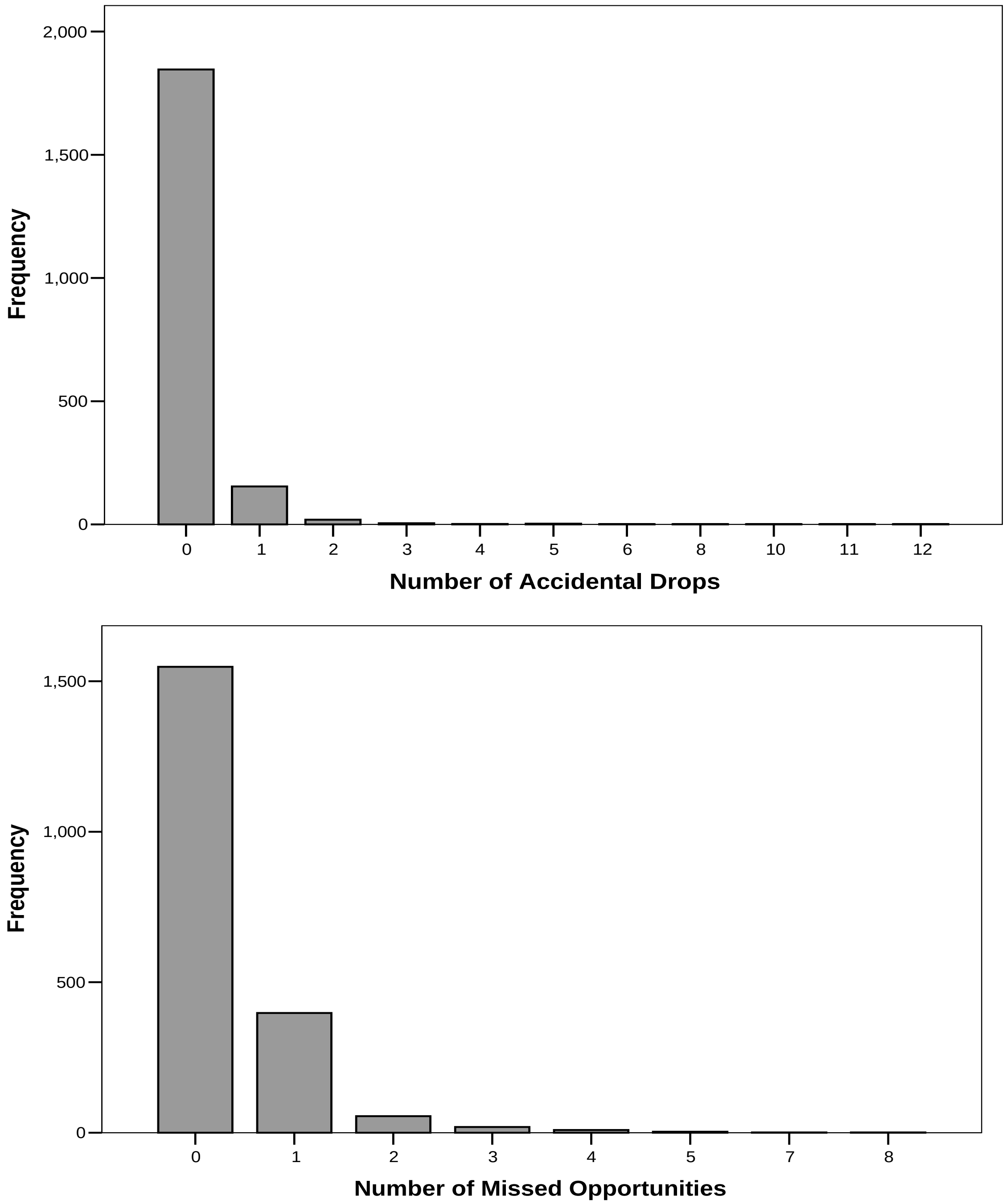


Figure 3.8 continued.

Analyses: Accuracy Measures

As demonstrated by Figure 3.8, the frequency for accuracy errors during the experimental task was generally low. In fact, the majority of participant trials were error free in terms of TX, AD and OND, and MDC errors occurred with a frequency of zero or 1 in most cases. Instead of evaluating these accuracy measures as continuous variables using linear regression, the accuracy measures were coded as dichotomous variables (e.g., 0 in those cases where no errors were committed and 1 in cases where 1 or more of that type of error was committed). Using these dichotomous variables, logistic regression models were generated to examine the impact of the predictor variables on the likelihood of committing an accuracy error. Using the same predictors as were considered with the efficiency measures, forward stepwise logistic regression, based on the likelihood ratio method, was used with the accuracy measures. As with the logistic regressions on the highlight times, the HL test was used to assess each model, along with the percentage of cases correctly classified.

Results: Clinically-acquired Ocular Factors

Table 3.15 summarizes the outcomes of the logistic regression models for OND, AD, TX, and MDC for the clinically-acquired predictors. All models demonstrated a failure to reject the null hypotheses of the HL test; the predicted values not significantly different from the observed dependent variables. In addition, the percentages reporting cases correctly classified by each model were at acceptable levels, especially considering the inherent variability of this population.

The coefficients (B and exp (B)), test statistics and significance for each efficiency measure are presented in Table 3.16. Figure 3.9 a-d reveals the magnitude of impact of each predictor variable on the model via exp (B), similar to the graphs used to explain the highlight time measures.

Table 3.15. Assessment of logistic regression models for accuracy measures.

Variable	HL Goodness of fit test	% Cases Correctly Classified
OND	p = 0.587	76.90%
AD	p = 0.561	90.80%
TX	p = 0.373	63.50%
MDC	p = 0.690	84.30%

Table 3.16. Summary of logistic regression outcomes for accuracy measures using clinically-acquired measures (***** indicates predictors excluded from the models).

OND				
Variables	**Coefficients (B)**	**Wald Statistic**	**p**	**Exp (B)**
Constant	0.085	0.095	0.758	1.367
SS	*****			
ISp	-0.506	54.316	<.001	0.603
AF	*****			
Column	*****			
Row	-0.704	57.210	<.001	0.495
Drop Location	-0.325	40.109	<.001	0.722
Trial #	-0.003	6.920	0.009	0.997
Age	*****			
SF-12 MCS	*****			
SF-12 PCS	*****			
Dexterity	*****			
NVA	0.694	8.455	0.004	2.002
CS	*****			
AMD Score	0.313	25.117	<.001	1.367
AMD Score* SS	*****			
AMD Score* ISp	*****			
AMD Score* AF	-0.350	19.896	<.001	0.704

Table 3.16. continued.

AD				
Variables	**Coefficients (B)**	**Wald Statistic**	**p**	**Exp (B)**
Constant	-4.901	18.763	<.001	0.007
SS	0.515	13.377	<.001	1.673
ISp	*****			
AF	*****			
Column	*****			
Row	*****			
Drop Location	*****			
Trial #	-0.004	4.776	0.029	0.996
Age	*****			
SF-12 MCS	*****			
SF-12 PCS	*****			
Dexterity	-0.285	64.821	<.001	0.752
NVA	-2.319	15.425	<.001	0.098
CS	0.187	18.149	<.001	1.206
AMD Score	0.573	11.046	0.001	1.774
AMD Score* SS	-0.168	5.918	0.015	0.845
AMD Score* ISp	*****			
AMD Score* AF	*****			

Table 3.16 continued.

TX				
Variables	**Coefficients (B)**	**Wald Statistic**	**p**	**Exp (B)**
Constant	-2.631	17.095	<.001	0.072
SS	*****			
ISp	*****			
AF	*****			
Column	*****			
Row	-0.358	25.471	<.001	0.699
Drop Location	*****			
Trial #	-0.002	4.239	0.040	0.998
Age	0.027	15.722	<.001	1.027
SF-12 MCS	*****			
SF-12 PCS	-0.028	24.971	<.001	0.972
Dexterity	0.061	7.934	0.005	1.063
NVA	1.426	36.286	<.001	4.160
CS	*****			
AMD Score	0.257	20.229	<.001	1.293
AMD Score* SS	*****			
AMD Score* ISp	*****			
AMD Score* AF	-0.231	12.431	<.001	0.794

Table 3.16. continued.

MDC				
Variables	**Coefficients (B)**	**Wald Statistic**	**p**	**Exp (B)**
Constant	-6.764	27.801	<.001	0.001
SS	*****			
ISp	*****			
AF	*****			
Column	-0.162	8.077	0.004	0.85
Row	*****			
Drop Location	-0.148	6.457	0.011	0.862
Trial #	*****			
Age	0.036	12.783	<.001	1.036
SF-12 MCS	*****			
SF-12 PCS	-0.020	7.034	0.008	0.98
Dexterity	-0.201	41.224	<.001	0.818
NVA	-3.698	54.602	<.001	0.025
CS	0.259	53.944	0.000	1.295
AMD Score	0.308	13.961	<.001	1.36
AMD Score* SS	*****			
AMD Score* ISp	*****			
AMD Score* AF	-0.317	16.795	<.001	0.728

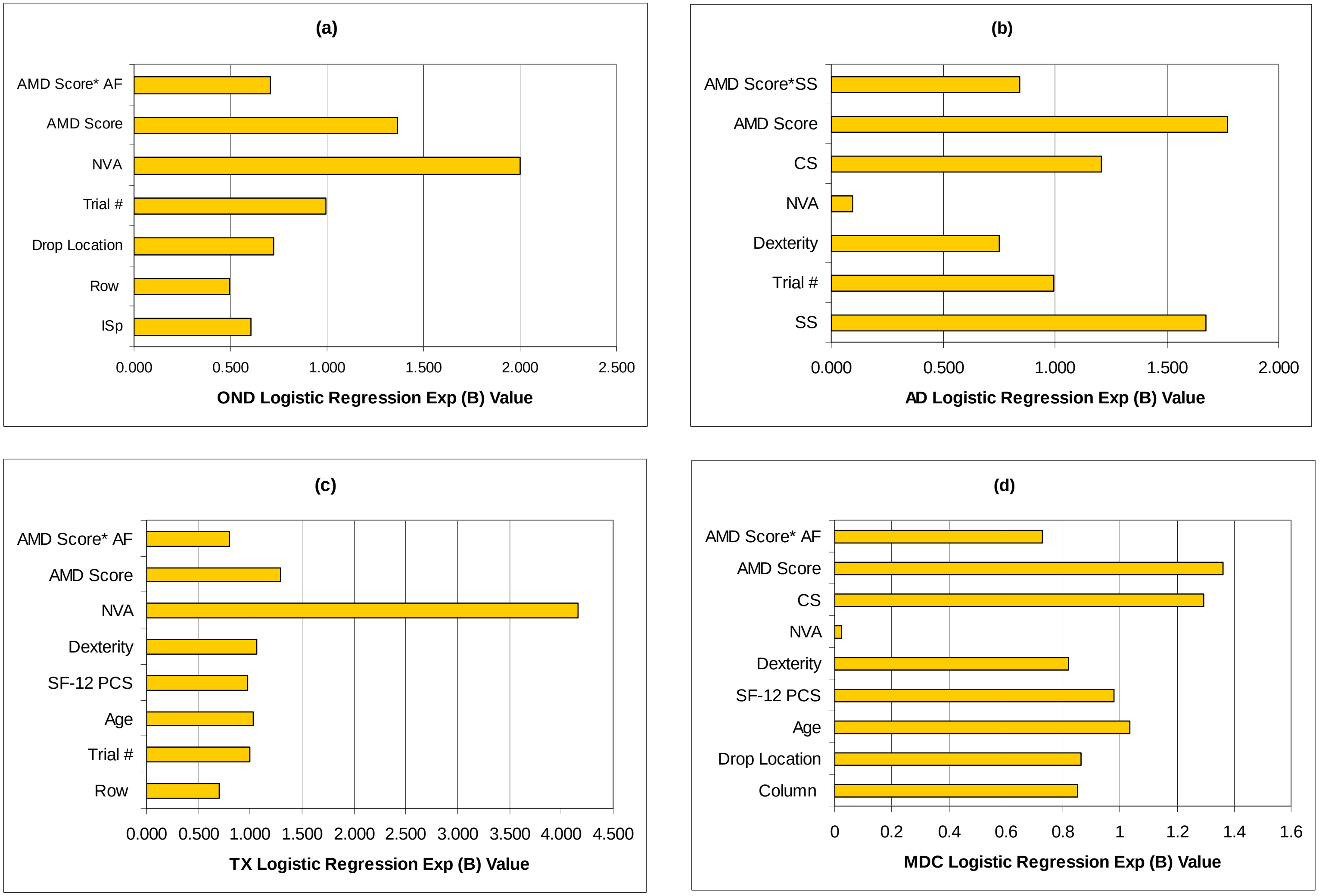


Figure 3.9. Illustration of relative impact of the predictor variables (clinically-acquired) on the likelihood of each error occurring at least once, Exp (B).

Results: Non-Clinically-Acquired Ocular Factors

Table 3.17 summarizes the assessments of the logistic regression models for OND, AD, TX, and MDC using the non-clinically-acquired ocular factors. All models resulted in a failure to reject the null hypotheses for the HL test; the predicted values were not significantly different from the observed dependent variables. In addition the percentages of cases correctly classified by each model were at acceptable levels. The coefficients (B and exp (B)), test statistics and significance for each efficiency measure are presented in Table 3.18. Figure 3.10 demonstrates the magnitude of impact of each predictor variable on the model via exp (B).

Table 3.17. Logistic regression model assessment for non-clinical ocular factors and accuracy measures.

Variable	HL Goodness of fit test	% Cases Correctly Classified
OND	p = 0.174	77.00%
AD	p = 0.554	90.80%
TX	p = 0.392	63.30%
MDC	p <.001	61.30%

Table 3.18. Summary of logistic regression outcomes for accuracy measures using non-clinically-acquired measures.

OND				
Variables	**Coefficients (B)**	**Wald Statistic**	**p**	**Exp (B)**
Constant	1.619	21.918	<.001	5.048
SS	*****			
ISp	-0.499	53.382	<.001	0.607
AF	-2.017	8.242	0.004	0.133
Column	*****			
Row	-0.683	54.953	<.001	0.505
Drop Location	-0.320	39.229	<.001	0.726
Trial #	-0.002	3.752	0.053	0.998
Age	*****			
SF-12 MCS	*****			
SF-12 PCS	-0.018	8.090	0.004	0.983
Dexterity	*****			
VFQ	*****			
VSAT	*****			
VFQ * SS	*****			
VFQ * ISp	*****			
VFQ * AF	0.021	6.640	0.010	1.021
VSAT * SS	*****			
VSAT * ISp	*****			
VSAT * AF	*****			

Table 3.18. continued.

AD				
Variables	**Coefficients (B)**	**Wald Statistic**	**p**	**Exp (B)**
Constant	-1.513	2.531	0.112	0.220
SS	*****			
ISp	*****			
AF	*****			
Column	*****			
Row	*****			
Drop Location	*****			
Trial #				
Age	0.034	11.303	0.001	1.035
SF-12 MCS	*****			
SF-12 PCS	*****			
Dexterity	*****			
VFQ	-0.042	60.930	<.001	0.959
VSAT	*****			
VFQ * SS	*****			
VFQ * ISp	*****			
VFQ * AF	*****			
VSAT * SS	*****			
VSAT * ISp	*****			
VSAT * AF	0.010	11.666	0.001	1.01

Table 3.18. continued.

TX				
Variables	**Coefficients (B)**	**Wald Statistic**	**p**	**Exp (B)**
Constant	-0.994	3.479	0.06	0.37
SS	*****			
ISp	*****			
AF	*****			
Column	*****			
Row	-0.236	10.227	0	0.79
Drop Location	*****			
Trial #	*****			
Age	0.014	7.856	0.01	1.014
SF-12 MCS	*****			
SF-12 PCS	-0.036	32.95	<.001	0.956
Dexterity	*****			
VFQ	0.019	17.889	<.001	1.019
VSAT	*****			
VFQ * SS	-0.004	22.833	<.001	0.996
VFQ * ISp	*****			
VFQ * AF	*****			
VSAT * SS	*****			
VSAT * ISp	*****			
VSAT * AF	*****			

Table 3.18. continued.

MDC				
Variables	**Coefficients (B)**	**Wald Statistic**	**p**	**Exp (B)**
Constant	3.456	110.360	<.001	31.680
SS	*****			
ISp	*****			
AF	*****			
Column	*****			
Row	*****			
Drop Location	*****			
Trial #	*****			
Age	*****			
SF-12 MCS	*****			
SF-12 PCS	-0.020	11.650	0.001	0.980
Dexterity	*****			
VFQ	-0.025	33.180	<.001	0.975
VSAT	*****			
VFQ * SS	*****			
VFQ * ISp	*****			
VFQ * AF	*****			
VSAT * SS	*****			
VSAT * ISp	*****			
VSAT * AF	*****			

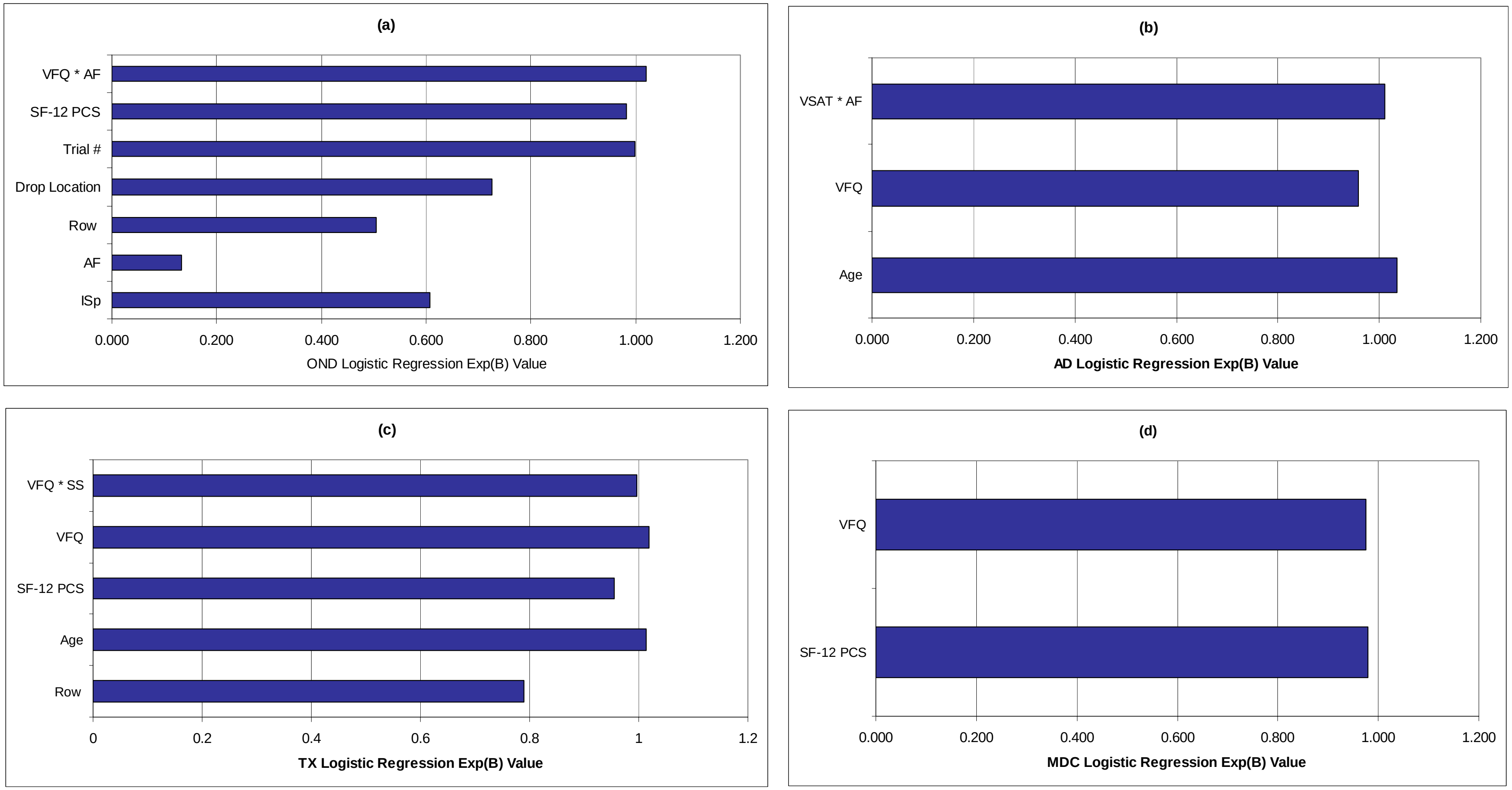


Figure 3.10. Illustration of the relative impact of the predictor variables (non-clinically-acquired) on the likelihood of each error occurring at least once

.

Accuracy Measure Outcome Summary

Clinically-acquired Ocular Factors

Outcome #1 Clinically-acquired Predictors

Akin to the models derived from the clinically-acquired visual factors for the efficiency measures, the models of accuracy were most heavily influenced by the ocular measures and interactions between the ocular measures and interface related independent variables. Both AMD Score and NVA were included as predictors for all four accuracy outcome measures, and CS was included as influential in the logistic regressions on two of the accuracy measures (AD and MDC).

AMD Severity. Under all four accuracy models, AMD Score was consistently the second most influential factor on the probability of error occurrence. Consistent with the models on efficiency, more severe cases of AMD influenced an increase in the likelihood of OND, AD, TX, and MDC models. This reveals a monotonic relationship between AMD and performance across all phases of the interaction, and makes interventions such as AF that significantly can lessen the bearing of AMD severity on performance, a priority.

Near Visual Acuity. NVA was the most influential predictor of the likelihood for each error to occur at least once. However, the direction of influence differs. As NVA worsened (from .1 to 1) the likelihood of OND and TX errors increased, while the probability for at least one AD or MCD was reduced in instances where near vision worsened. In the model the impact of diminished NVA on performance was not monotonic. As near vision declined, the likelihood for error in the 'drop' phase of the task was prone to increase for on OND and TX. This identified possible issues in the precision in path approach for NVA deficits. However, the model indicated that declines in NVA imposed a lower probability for errors of commission in the movement of the card (via MDC and AD, in addition to DD). This implies that the participants NVA created obstacles for particular phases of the task, especially the drop, the efficient use of the stylus to move from point A to point B. It also implies that the speed accuracy trade-off is limited the icon movement portion of the task, in terms of contrast sensitivity, since the measures of drop efficacy (OND) both decreased in the presence of declines in NVA.

Contrast Sensitivity. CS was only influential in 2 models (in contrast to its inclusion in all of the efficiency models), AD and MDC. Both measures detail the underlying movement and placement of the icon in the drop pile. CS scores that were higher (or improved) imposed an increase in the probability of both AD and MDC. This reflects the same unexpected result that emerged in the inclusion of NVA for these two measures.

As visual function worsened (based on decreasing CS score and/or NVA values increasing to 1), the models reflected a decrease in the probability for AD and MDC. This is particularly interesting in light of the trend for functional declines in NVA and CS to influence shorter distances, but slower movement times, effectively a speed accuracy trade-off on the movement portion of the task. Individuals with poor CS and/or NVA may take more time in moving the icon, as a result of their limited bandwidth of visual perception, with the residual impact of increased accuracy.

Outcome #2 Controlled Interface Components

Set Size and Icon Spacing. In the clinical models of error likelihood, SS and ISp demonstrated an effect across all participants, but not across all of the accuracy measures. An increase in SS generated a higher likelihood of AD errors. The longer distance to travel with the icon and the presence of distracters while attempting to visually track and move the icon may be a potential source of this effect.

ISp only affected the OND measures; as ISP increased, there was a decrease in the probability of OND occurrence. This is consistent with the effect observed in the models of TTHT, where increased ISp influenced a faster release of the icon once it was in correct position for the drop.

Outcome #3 Interface and Visual Factor Interactions

*AMD*AF.* Not unlike the efficiency models, the impact of the auditory feedback on task accuracy was observed through its interaction with the AMD Score, disease severity rating. The impact of AF was not realized across all participants for the models, but instead, only when the diagnosis of the individual was positive for AMD. Furthermore, the extent to which AF influenced the model was coupled with the level of disease severity: The more severe the AMD diagnosis, the lower the probability that the participants were likely to experience OND, TX, or MDC. The impact of the interaction may offset or cancel the negative influence of disease in the model, demonstrating that AF provides helpful cues even in the context of distracters and in the absence of truly ‘intelligent’ AF with different sounds mapped to specific icons on the display.

*AMD*SS.* The interaction between AMD Score and SS was included in the model of the probability of AD error occurrence. As in the efficiency measures for visual search, participants, with higher severity of the disease were not negatively affected by increases in the number of distracter icons on the screen, as the controls. This is attributable to the same justification that was provided in the visual search model. As AMD increases in severity, there is a higher likelihood for distortions, aberrations, and interruptions to the visual field. This implies that despite the fact that some the experimental trials included greater numbers of icons; there may not have been a noticeable increase within the individual’s useful visual field. That said, their strategy for scanning the display might be the same under each condition of SS, because the number of icons perceived at the start of the task effectively did not change at the onset of display presentation.

Outcome #4 Personal Factors

Age. As participant age increased in the models of interaction accuracy, participants were more likely to experience TX, AD, and MDC. This poses another instance of speed accuracy trade-off in the interaction, as increases in age were found to influence faster TT, VST and MT. Perhaps the older participants were overconfident in their ability to move the cards, and the effects of age on hand-eye coordination were driving less accurate task performance.

Dexterity. For the clinical model, increases in dexterity were influential in decreasing the probability on three of the four accuracy measures. Not surprising, as dexterity improved in the model, the likelihood for MDC, TX and AD diminished. These results are not unexpected since the accuracy measures truly captured the efficacy of users' manipulation of the stylus in coordination with their visual ability.

SF-12 PCS. As the self reported assessment of physical health, as measured by the SF-12, improved, the likelihood of TX and MDC errors were curbed. This is consistent with the outcomes based on dexterity and Age, and logical with the required manipulation. However, in the clinical models of efficiency, PCS was shown to actually increase the time metrics (decreasing efficiency on the task). Therefore, this revealed another speed accuracy trade-off: those individuals who rated their physical health as higher were more likely to interact with more accuracy on the task, but at the cost of longer task and visual search times.

Outcome #5 Other interface features

Trial Number. There was a very small learning effect included in the clinical models of the accuracy measures – the least influential factor of those predictors included in the models of OND and TX.

Row and Column of target card icon and Drop Pile Location

Those icons placed in rows lower on the display influenced the model towards lower probabilities for OND and TX. Columns further to the right of the display increased the likelihood of MDC. This is interesting from the perspective that the location is not only imposing longer or shorter times on the task, which is an intuitive result. Drop Locations lower on the display also influenced decreases in OND, and MDC, consistent with the results of the models for target icon row. The accuracy measures inform that certain areas of the display were more amenable to the interactions required of the participants with the stylus and icons. What's more, this result promotes the future consideration of placement of interaction on the display – even in the context of small displays.

Non-Clinically-Acquired Ocular Factors

Outcome#1: Differences between Clinical and Non-Clinical Models

Just as with the efficiency measures there were some differences that emerged between the clinical and non-clinical analyses on accuracy, highlighted in Table 3.19. However, in the accuracy models, none of the emergent results were contradictory to what emerged from the clinically-based models. Notably, the predictors in the non-clinical models were fewer in number, especially in the case of AD and MDC.

Table 3.19. Differences between clinical and non-clinical models derived for the accuracy measures.

Model	Consistent	Unique to Clinical Models	Unique to Non- Clinical Models
p(OND ≥ 1)	ISp Row Trial # Drop Location	AMD Score NVA AMD*AF	AF VFQ*AF SF-12 MCS
p(AD ≥ 1)	NA	CS NVA SS AMD*SS Dexterity Trial#	VFQ VSAT*AF Age
p(TX ≥ 1)	Age SF-12 PCS Row	AMD Score NVA AMD*AF Dexterity Trial#	VFQ VFQ*SS
p(MDC > 1)	SF-12 PCS	AMD Score CS NVA AMD*AF Age Dexterity Column Drop Location	VFQ

Outcome #2 Non-Clinically-acquired Factors

Self-Perceived Visual Functioning. In terms of main effects for the non-clinical visual factors, VFQ overall score was the only main effect observed. Increases in VFQ decreased the likelihood of AD and MDC but increased the likelihood for one or more TX. In practical terms, this characterizes the efficacy of the approach to the drop pile with the card icon. According to the model, individuals with better perceptions of their visual function were more accurate in their approach to the drop pile, but crossed the task axis more often before dropping the card icon into the pile.

Outcome #3 Controlled Interface Factors

Inter-Icon Spacing. A lower likelihood for OND errors was imposed in the model based on increases in inter icon spacing. The further the icons were spaced, the more likely the participants were to release the card into the pile when in appropriate position for an acceptable drop. This result was consistent with the role of ISp in the clinically-acquired model.

Auditory Feedback. The presence of auditory feedback imposed a large influence on decreased OND. The likelihood for and OND event decreased substantially in the presence of the supplemental sound. This suggests that ISP and AF are potentially positive influences on the drop portion of the task across all participants.

Outcome #4 Interactions between Visual and Interface Factors

*VFQ*SS.* The increasing interaction between perceived visual function and set size was observed to impose decreases in the likelihood for TX errors. The interaction term, however, did not impose great influence on reducing the probability TX in the model, especially relative to the other predictors in the model. Therefore, while it did bear some influence, it did not drastically affect the TX as much as other factors (PCS, ROWS). More importantly, however, this interaction demonstrates that individuals with better perceived vision are more accurate in their manipulations of icons in the presence of more distracters. Increasing SS and holding everything else constant generated, in the model, a wider performance differential between individuals who differ based on their levels perceived visual function.

*VFQ*AF and VSAT*AF.* The likelihood for OND increased in the presence of increasing VFQ*AF and the likelihood for AD increased in those cases with increasing VSAT*AF. For OND, because AF decreased the likelihood across all participants to a greater degree than the VFQ*AF term increased the likelihood, reveals that the AF was not hindering the performance of the visually healthier participants. Instead, the self perceived visually healthy participants in this model were not benefiting from the feedback as much as those with lower perceived visual function.

The VSAT*AF predictor increased the likelihood for AD errors to occur. This term indicates that there was an increase in AD across all participants in the presence of the supplemental non-visual cue, but only proportional to the participants' visual attention score. The better the VSAT score in the presence of AF, the more likely the AD was to occur. With all other terms constant, a decrease in VSAT in the presence of AF would lessen the negative influence of AF on the AD measure. That said, this implies that AF can be detrimental to performance, especially in cases of where the non-clinical visual factors were not indicative of visual dysfunction.

Outcome #5 Personal Factors

SF-12 PCS. Improvements in PCS score (increasing PCS) decreased the likelihood for OND, TX and MDC. This suggests that the manipulation of the icons with the stylus were easier for those who rated their physical health as higher. Furthermore, this result is consistent with the clinical models for TX and MDC. It is intuitively reasonable because a number of physical functions have to work in coordination during HCI (i.e. the visual sensory, motor skills, hand-eye coordination and also endurance and continued focus over the course of 162 trials). Those individuals who experience decrements to their physical health may have more difficulty in accurately completing the task.

Age. Increases in age were observed to influence increased probabilities for both AD and TX, and is consistent with the effects of the SF-12 PCS and the non-clinical models of the efficiency measures, and the clinical models of accuracy. This is not unexpected as aging is associated with a variety of impediments to vision, coordination and fatigue that may not be captured wholly by other predictors in the models. Pervious studies on HCI have observed a similar influence on aging for a both a larger participant age range (Edwards, Barnard et al. 2005). The results form the current study indicated the relevance of the differences between the different age groups within the general older adult population.

Outcome #6 Other interface features

Trial #. There was a small learning effect modeled for the OND errors. Later trials influenced a small decrease in OND. While accounted for in the model OND, it is of importance that the visual function variables and controlled interface features greatly override the influence of learning over the course of the trials. This is also demonstrated via the other three accuracy models' exclusion of Trial # as a predictor.

Row of target card icon & Drop Pile Location. The column was not included as a predictor for any of the accuracy measures in the non-clinical models. However, the influence card icon rows and drop piles lower on the screen improving accuracy of the modeled interaction components. This was observed for drop location for OND error, and for card icon row location on both OND and TX.

3.4 Information Processing

General Summary

Passable eye tracking data were gathered form 11 of the 13 participants for the handheld computer component of the investigation. While data were recorded on the other two participants, its fidelity was questionable due to inability to capture and hold a consistent image of the pupil and corneal reflections. Of the 11 data sets, 3 were from controls, and the other 8 were from participants that were diagnosed with AMD. The distribution of the eye movement summary measures, fixation duration, saccade duration, and saccade to fixation duration ratio, are summarized in Table 3.20. While the number of fixations actually recorded by the eye tracking system deviated between participants based on the quality of eye tracking and their length of time on task, other non-frequency/count measures provided insight into the implicit aspects of information processes.

The relative distributions of the saccade and fixation duration for all participants are shown in Figure 3.11. This graphic demonstrates the how there was a higher amount variability in the saccade duration metric for these participants compared with their fixation duration. Figure 3.12 illustrates the ranges observed in the participant saccade to fixation duration ratio. Those bars that extend above the value 1 are indicative of the participants who spent more time in pursuit of items to fixate (a.k.a. saccades) than in actual fixations (or processing information).

Pupillary response provides an indication of the mental workload encountered by the participant during the task. A time based measure, pupillary response, as calculated by Backs and Walrath (Backs and Walrath 1992) was used, and reported in mm change. Figure 3.13 illustrates the change in pupil diameter, or pupillary response, through a stacked plot of the minimum and maximum value for each participant.

Table 3.20. Summary of physiological measures of information processing.

Physiological Measures	
Fixation Duration (seconds)	Mean (Std. Error): .366 (.003) Median: .25 Minimum: .033 Maximum: 5.138
Saccade Duration (seconds)	Mean (Std. Error): .334 (.0285) Median: .0330 Minimum: <.0001 ms Maximum: 243.544
Saccade to Fixation Duration Ratio	Mean (Std. Error): 3.013 (1.820) Median: .6338 Minimum: .15 Maximum: 20.52
Pupillary Response (mm)	Mean (Std. Error): 1.944 (.3202) Median: 1.615 Minimum: .715 Maximum: 4.511

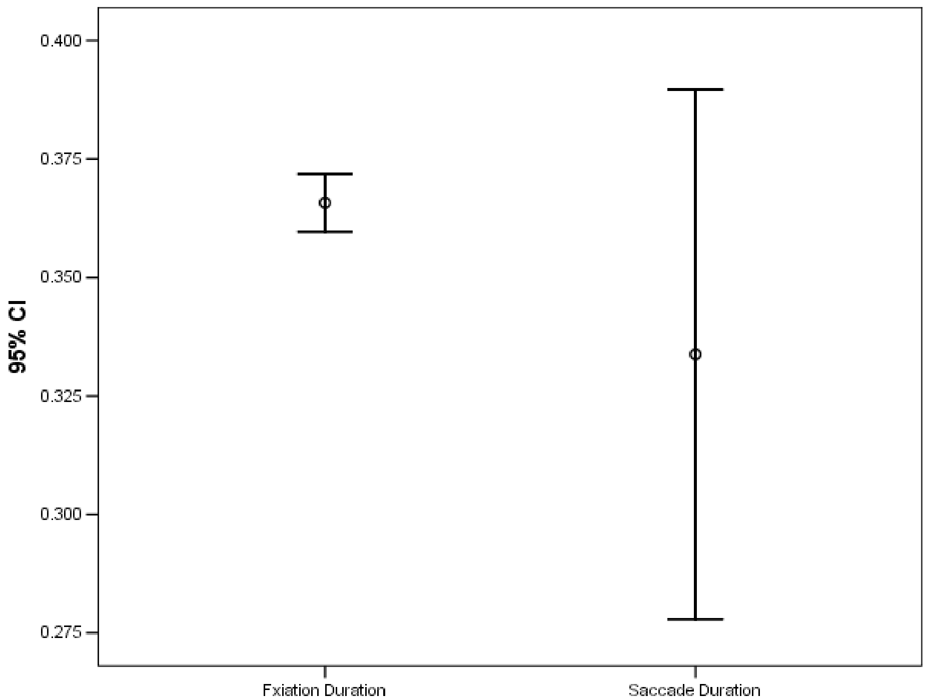


Figure 3.11. Distribution of fixation duration and saccade duration for participants.

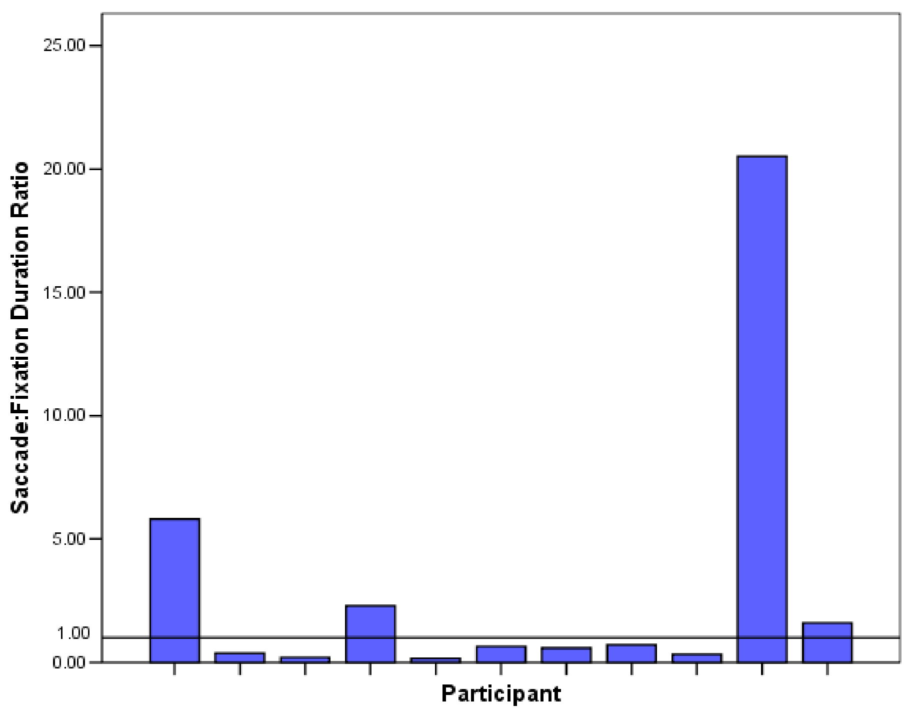


Figure 3.12. Saccade to fixation duration ratio, the reference line at 1.00 provides an indication of the number of participants whose time spent searching surpassed the amount of time in fixating on the display or processing information.

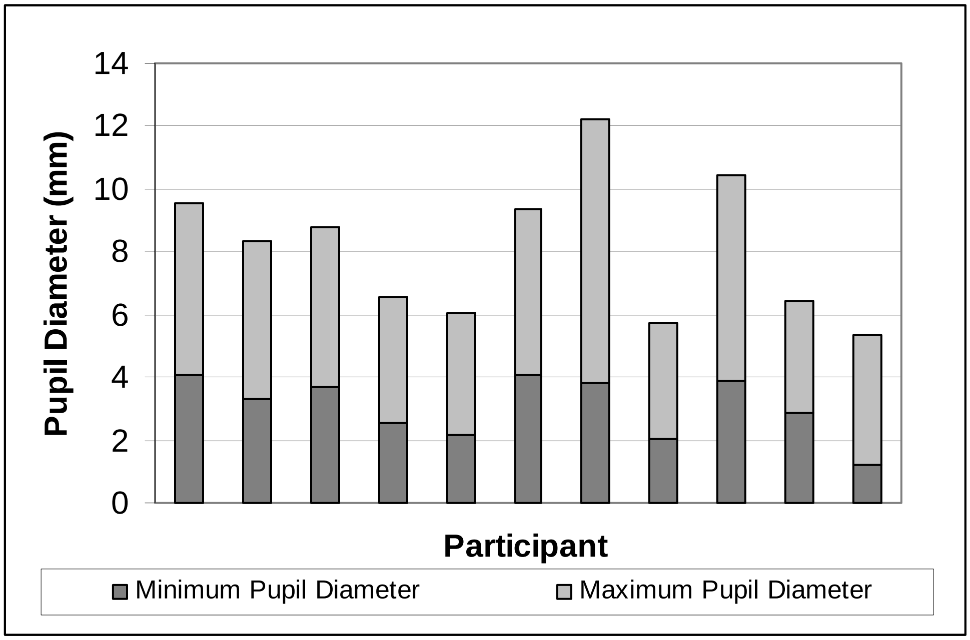


Figure 3.13. Illustration of the maximum and minimum pupillary response values for each participant.

Analyses: Information Processing

The eye movement and pupillary response data, unlike the other dependent measures, were not captured at all levels of the independent interface factors (SS, ISp, and AF). Instead, the eye data summarized the entire task for each participant. The limited number of cases therefore triggers a lack of diagnosticity of the regression analyses to detect relationships. In addition, due to the small sample size and, the eye data (saccade and fixation) did not meet the prerequisites to use parametric statistics, even after attempted variable transformations. These measures are therefore considered to be global indices of the information processing experienced during task interaction.

Instead of regression models, non-parametric statistical comparisons were made, using the Mann-Whitney test (Field 2000; Pallant 2003). Two groups were created for each eye tracking metric, based on the midpoint (50^{th} percentile) of the data from the 11 participants (see Miller 1987 for an example of this approach in the psychology literature). These groups are summarized by Table 3.21 and Figure 3.14. Mann-Whitney comparisons showed that the two groups differed significantly for each eye tracking measure from which they were derived. (e.g., Group 1 for Mean Fixation Duration had significantly lower Mean Fixation Durations than Group 2), at $p \leq .006$. For each measure comparisons between the two groups were made on all the clinically-based ocular measures, non-clinically based ocular factors and participant-based factors. While these separate comparisons are not as diagnostic in predicting the association of the predictor variables with each other and the eye metrics, they can yield a useful starting point in understanding the true nature of the interactions, and provide a baseline for future studies. Particularly, the emergent differences can provide additional insight into the results of the regressions.

Table 3.21. Summary of Eye movement metric group classification.

	Fixation Duration (Mean)	Saccade Duration (Mean)	Saccade to Fixation Ratio	Pupillary Change
50th percentile	0.375 sec	0.1626 sec	0.634	1.615 mm
Group 1	< .3752 n = 5	≤ .1616 n = 6	≤ .6338 n = 6	< 1.615 n = 5
Group 2	≥ .3752 n = 6	> .1616 n = 5	> .6338 n = 5	≥ 1.615 n = 6
Mann-Whitney Comparison	Z = -2.739 p = .006	Z = -2.739 p = .006	Z = -2.739 p = .006	Z = -2.739 p = .006

< x means the scores in that group are all less than x; ≤ means the scores in that group are all less than or equal to x

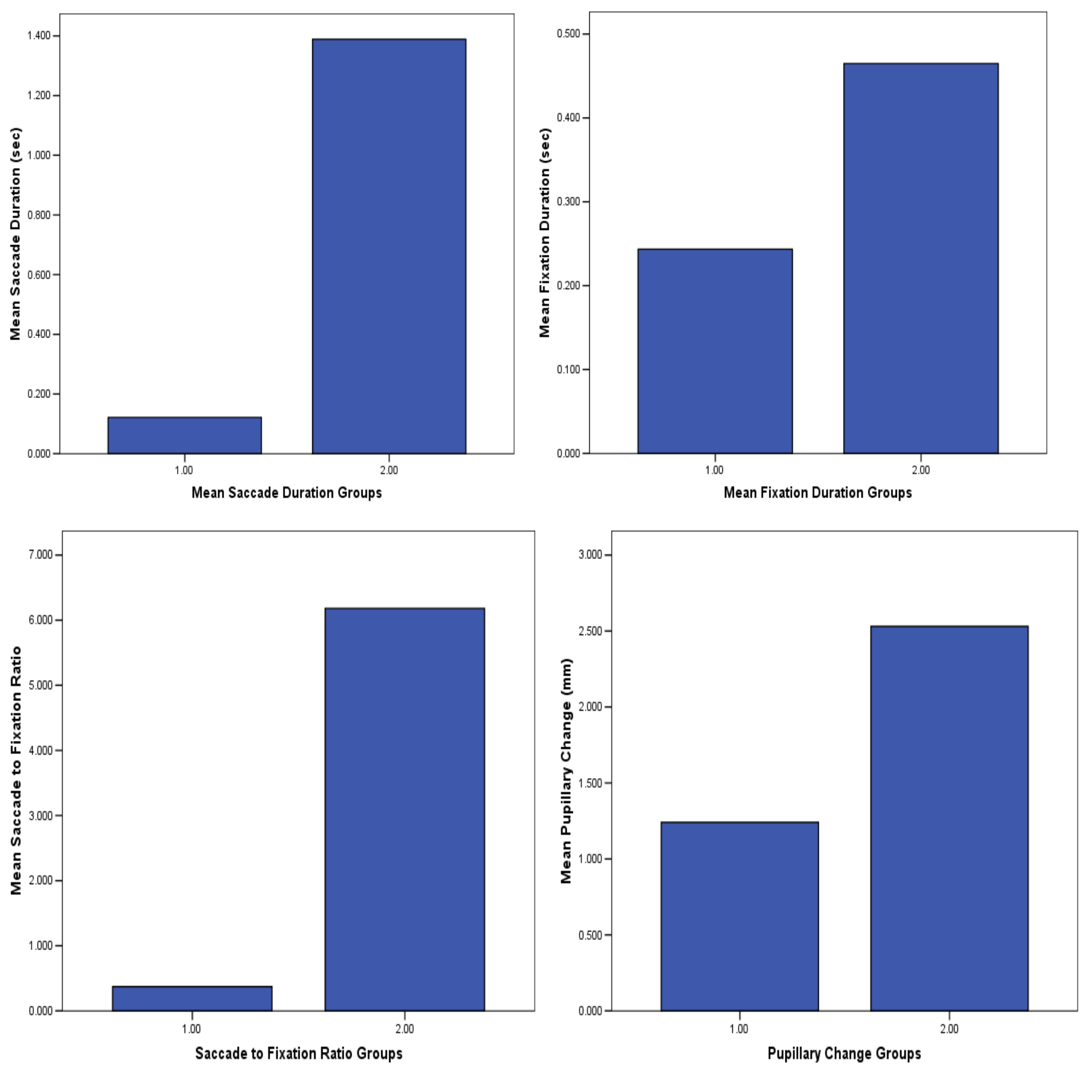


Figure 3.14. Summary of groups derived from eye tracking measures.

Results: Information Processing

As stated, Mann-Whitney tests were used to compare the differences in personal factors, clinically-acquired factors and non-clinically-acquired factors between the two groups on each eye tracking metric. Interestingly, the significant differences between the groups for these factors were limited, but insightful.

Contrast sensitivity (CS) was significantly different between the groups based on fixation duration. Specifically, the group with the longer mean fixation duration (Group 2) demonstrated a significantly lower, or worse score for CS (Z = -.034; p = .034). For mean saccade duration, those participants grouped as having longer saccade duration revealed a trend with significantly lower *mental health scores,* as rated by the SF-12 MCS (Z = -2.001; p = .045). SF-12 MCS was also significantly different between the two groups in the analysis of the saccade to fixation duration ratio.

For those grouped as having significantly higher saccades to fixation ratios, scores on the *SF-12 MCS* were significantly lower (Z = -2.01, p = .045) than those participants in the group with lower saccade to fixation ratios. In terms of pupillary change over the course of the task, the group with a larger pupillary response had significantly lower, or worse, scores on the assessment of *CS*. It is interest that none of the non-clinically-acquired ocular factors showed signs of significant deviation between the two groups within any of the eye tracking metrics.

Figures 3.15-3.17 illustrate the differences between the groups, for each eye tracking metric, grouped by personal factors, clinically-acquired, and non-clinically-acquired ocular factors. The only personal factor to emerge as significantly different was SF-12 MCS for both saccade duration and saccade to fixation ratio, while the only clinical ocular factor to significantly differ between the groups was *CS* for both mean saccade duration and pupillary response.

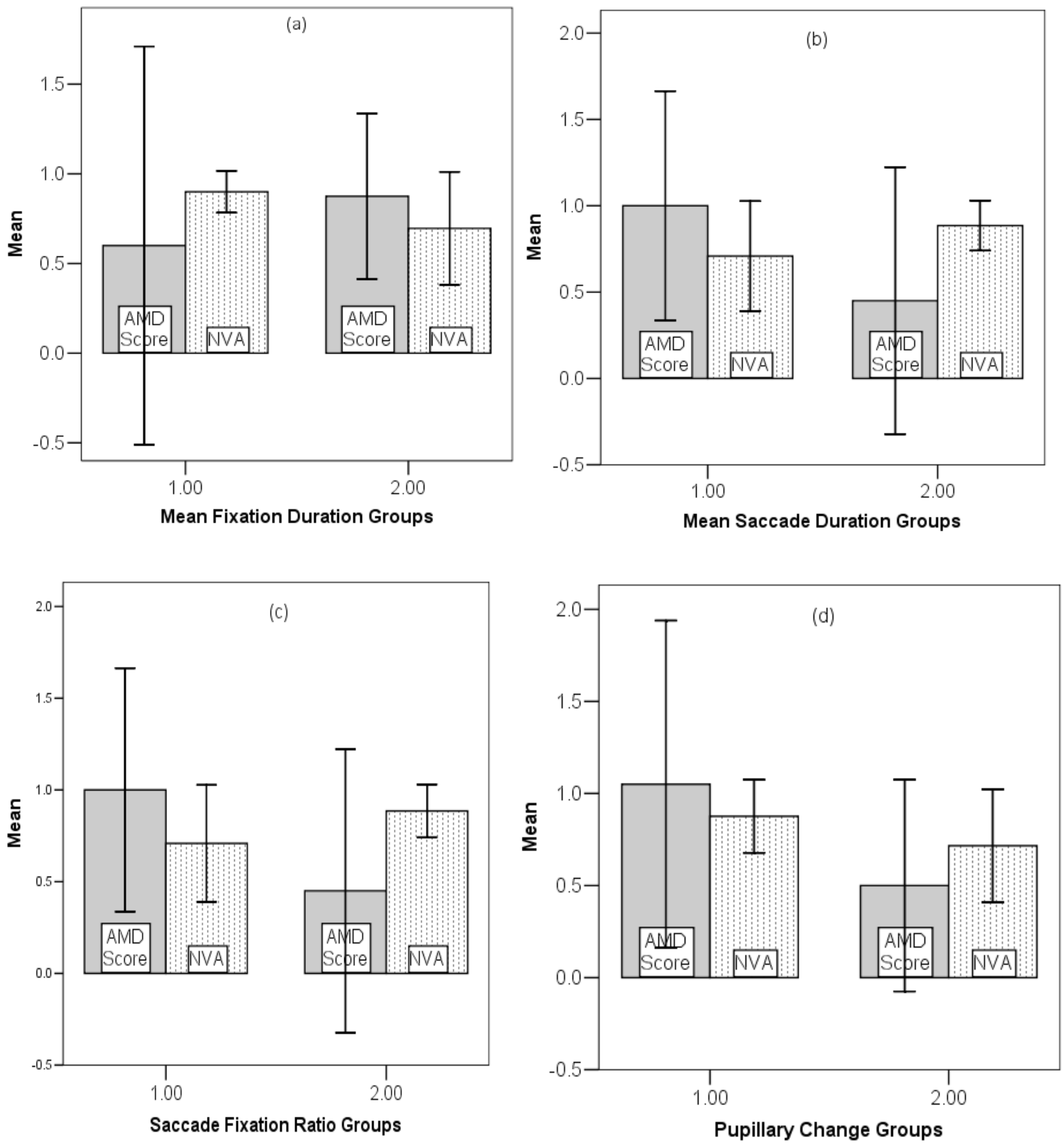


Figure 3.15. 95% CI for clinically-acquired ocular factors according to eye tracking metric and group; there were no significant differences between groups for these factors.

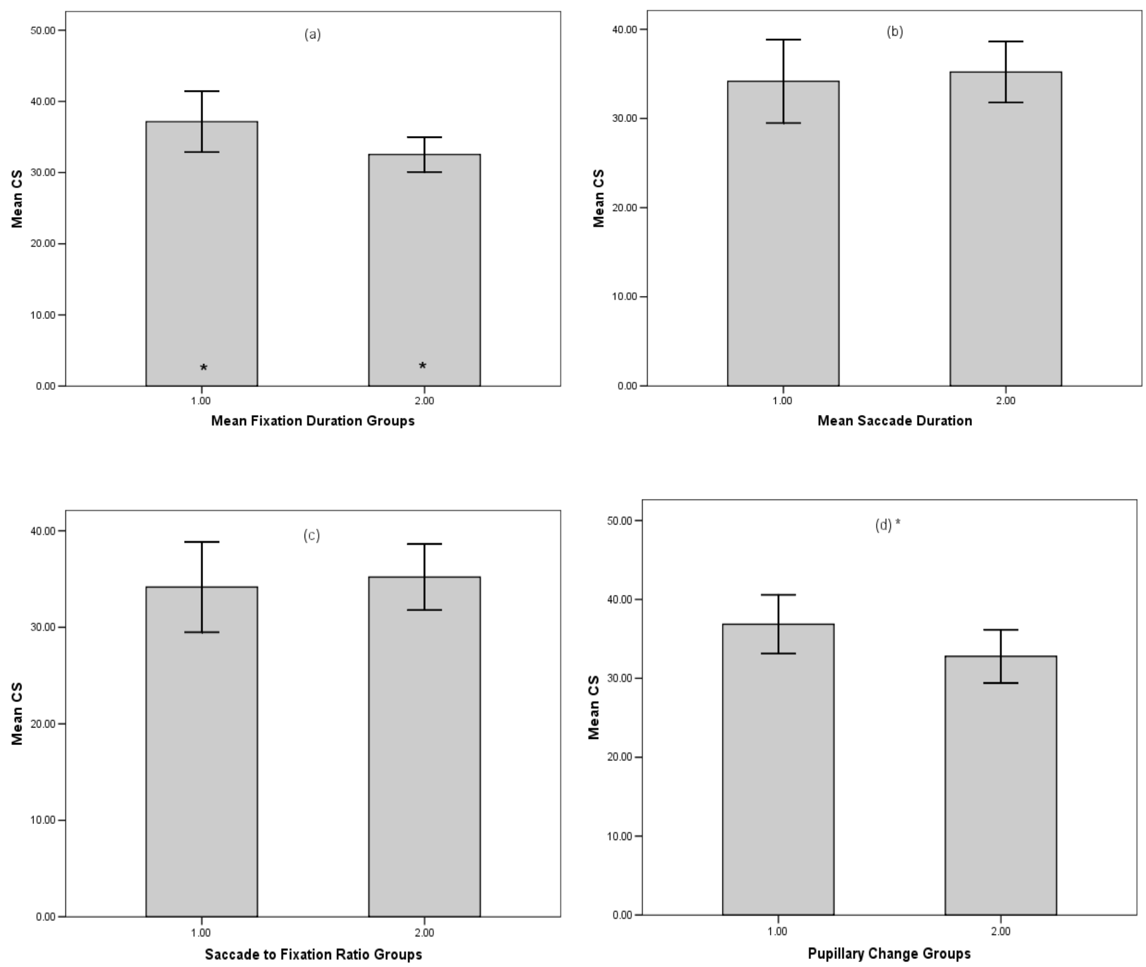
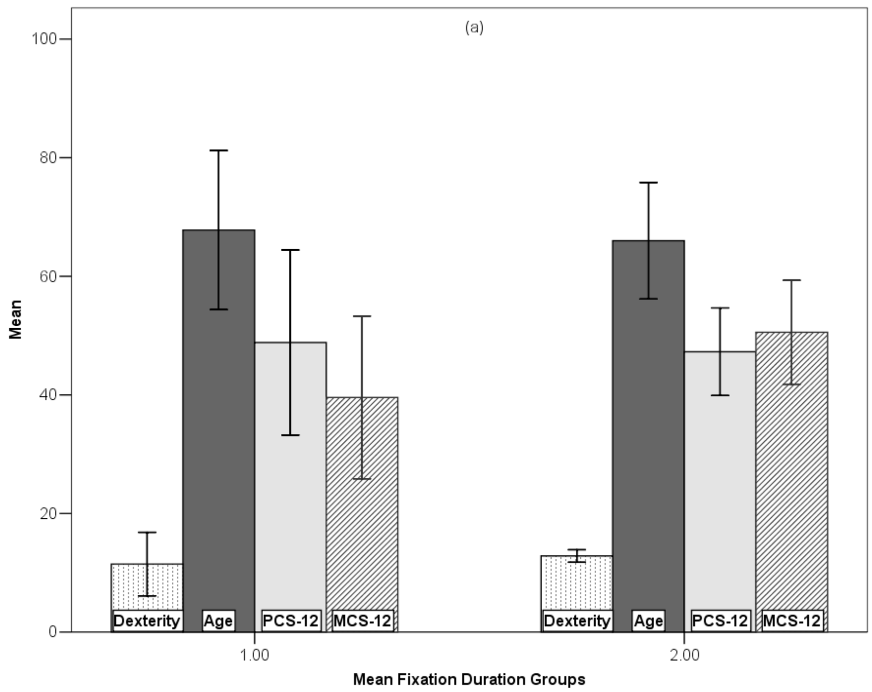


Figure 3.16. 95% CI for another clinically-acquired visual factor, contrast sensitivity, according to eye tracking metric and group; * designates factors for which there was a significant difference between the groups (p <.05).



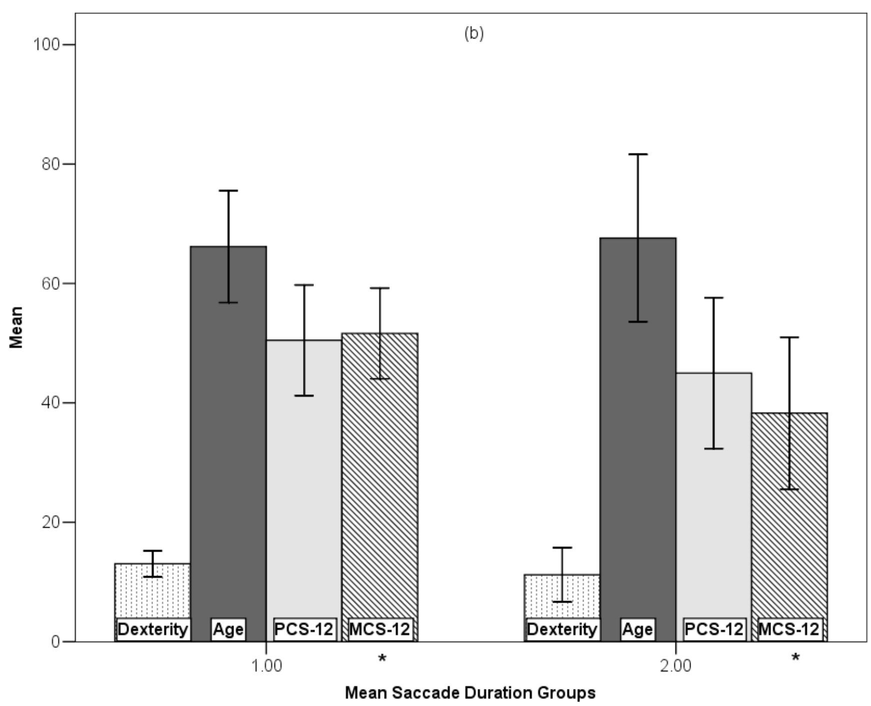


Figure 3.17. ((a)-(d)) 95% CI for personal factors according to eye tracking metric and group; there were no significant differences between the groups for these factors.

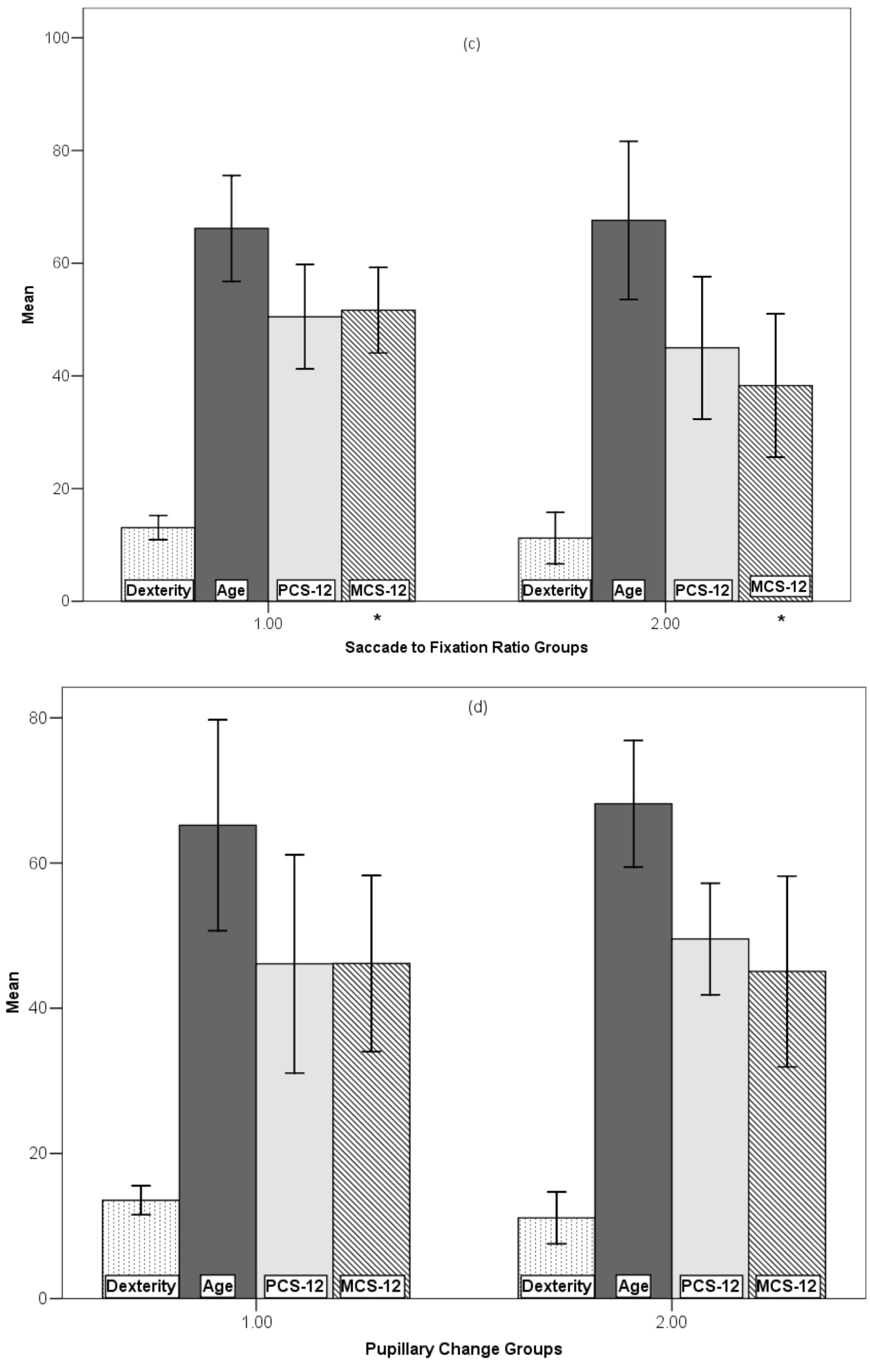


Figure 3.17. continued.

Information Processing Outcomes

<u>Outcome #1 Contrast Sensitivity and Information Processing</u>

Mean Fixation Duration. The group with significantly longer mean fixation durations also had significantly lower scores in the clinical assessments of their contrast sensitivity. Fixation duration is commonly used a measure of information processing and workload. Mean fixation durations that are longer imply more time spent in the interpretation of the visual cue or matching the visual stimulus to an internalized representation. In general, the longer the fixation, the more difficulty the participant has retrieving the internalize representation Goldberg and Kotval (1999).

This is relevant when considering the effect of decreased contrast sensitivity. Decreases in contrast sensitivity would impose difficulties in perceiving the card from the background, and more importantly could encumber the discrimination of the visual icons and numbers from the background. Therefore, the image they perceive poorly matches what is stored in their internal memory, and fixations are longer as they resolve the differences between these to assure they choose the correct icon. The models of TT and VST, which showed faster rates in the presence of improved CS, support this.

Pupillary Change. A similar effect of CS was observed in the analyses of pupillary change. The group with significantly greater deviation in their pupillary response during the task also had significantly worse CS. The larger the deviation in pupillary response suggests a higher level of mental workload by participants in this group. Based on this interpretation, participants with the higher level of mental workload, as measured by pupillary response over the course of the task had significantly worse CS the other participants. This lends explanation to the observation of poor CS influencing longer VST, TT, MT, but decreased DD, AD and MDC. These participants exerted a greater amount of effort over the entire course of the task, and in the instance of icon movement to the drop pile, they are more effective, but have to work harder to track the icon movement across the display (e.g. through more fixations, and more mental concentration/workload). The processing of the visual information was quantifiably more involved for those individuals with lower measurable function in contrast sensitivity.

Outcome #2 SF12-MCS and Information Processing

Mean Saccade Duration. The participants in the mean saccade group with longer mean saccade durations also demonstrated significantly lower scores on the SF-12 MCS, or lower levels of mental health. Mental health is an important consideration in HCI for older adults, as declines in working memory are common. The increased saccade duration suggests that the visual perception of the interface was less directed and less efficient, which is closely linked with an individual's mental ability.

Saccade to Fixation Duration Ratio. The ratio of the total saccade duration to total fixation duration is a measure of overall efficiency in the task. While saccade duration is also indicative of efficiency issues, it could be the case that an individual's visual perception of the display is slow over all, and that the mean fixation duration is also longer. Saccade to fixation ratio measures the amount of time spent searching the interface, as opposed to time spent processing information. This measure also confirmed the effects of lower rated mental health being associated with less efficient visual perception and information processing mechanism. The group with the higher saccade to fixation ratio had significantly lower mental health scores. Essentially, this suggests that lower mental health inhibited the ability to efficiently visually scan the display during interaction, causing a higher amount of workload in the acquisition of visual information and coordination with the mental processes.

3.5 Consolidation of Model

The considerable number of results presented in the analyses of task efficiency, accuracy, information processing, for both clinically and non-clinically-acquired visual factors, personal factors, and task factors generated a large number of interesting results. However, the global interpretation is needed to effectively glean the most compelling results. Tables 3.22-3.25 provide the opportunity to evaluate the comprehensive set of results, and draw out the most significant patterns. A separate table summarizes for each class of predictor variables and the interaction terms. The results within each are organized by the various dependent variables that were captured. When appropriate, the source of the result is noted; whether it was generated from the clinical models or the non-clinical models. For each predictor, the tables demonstrate the general relationship with the outcome measures based on increases in the predictor value with the practical meaning of ‘increase’ provided for the context of each predictor variable. Lastly, the table includes an indication of the directional impact of the increase in each predictor value on the outcome variable. *Increase* notes those outcomes observed to increase in value with an increase of the predictor, and *Decrease* signifies those outcomes that were smaller in the presence of increases in the predictor variable. These tables provide a useful way to extract the general trends emergent from theses analyses, while the bar graphs in each section are more useful in quantifying the magnitude of the impact.

Table 3.22. Summary of outcomes attributed to the visual factors.

Clinical	**Predictor: AMD Score** Increased disease severity level (worse)	
	Increasing	TT
		VST
		MT
		DD
		TTHT
		OND
		AD
		TX
		MDC
	Predictor: CS Score Higher, improved contrast sensitivity	
	Increasing	DD
		AD
		MDC
	Decreasing	TT
		VST
		MT
		Pupillary Change
		Fixation Duration
	Predictor: NVA Score Decrease in visual acuity, worse vision	
	Increasing	TT
		MT
		TTHT
		OND
		TX
	Decreasing	DD
		AD
		MDC
Non-Clinical	**Predictor: VFQ Overall** Improved perception of visual function	
	Decreasing	DD
		MDC
		AD
		TX
	Predictor: VSAT Percentile Higher percentile for visual attention (improved)	
	Decreasing	MT
		VST
		TT

Table 3.23. Summary of outcomes based on the independent task related factors; the source of the outcome, clinical or non-clinical models, is identified with a respective checkmark in the appropriate column.

Predictor: AF AF absent - AF present		**Clinical**	**Non-Clinical**
Decreasing	MT		✓
	TT		✓
	OND		✓
	TTHT	✓	
Predictor: SS Increasing the number of icons		**Clinical**	**Non-Clinical**
Increasing	TT	✓	✓
	VST	✓	✓
	AD	✓	✓
Predictor: ISp Increasing the inter-icon spacing		**Clinical**	**Non-Clinical**
Increasing	DD	✓	✓
Decreasing	TTHT	✓	
	OD	✓	✓

Table 3.24. Summary of the interactions between the visual factors and the interface independent variables.

Clinical	**Predictor: AF*AMD Score**	
	Decreasing	MT
		DD
		TTHT
		OND
		TX
		MDC
	Predictor: SS*AMD Score	
	Decreasing	TT
		VST
		AD
	Predictor: ISp*AMD Score	
	NONE	

Cli	**Predictor: AF*VFQ Overall**	
	Increasing	TT
		VST
		MT
		OND
	Predictor: SS*VFQ Overall	
	Decreasing	TX
	Predictor: ISp*VFQ Overall	
	Decreasing	TTHT
	Predictor: AF*VSAT Percentile	
	Increasing	VST
		p(AD≥1)
	Predictor: SS*VSAT Percentile	
	Increasing	TT
	Predictor: ISp*VSAT Percentile	
	Increasing	TT
		VST

Table 3.25. Summary of the impact of personal factors and their impact on the dependent variables; check marks indicate which model reflects the given results.

Predictor: Age Older age		**Clinical**	**Non-Clinical**
Increasing	TT		✓
	VST		✓
	MT		✓
	DD		✓
	MDC	✓	
	TX	✓	✓
	AD		✓
Decreasing	TT	✓	
	VST	✓	
	MT	✓	
Predictor: Dexterity Improved manual dexterity		**Clinical**	**Non-Clinical**
Increasing	MT		✓
Decreasing	TT	✓	✓
	VST	✓	✓
	DD	✓	
	TTHT	✓	
	AD	✓	
	TX	✓	
	MDC	✓	
Predictor: SF-12 PCS Improved rating of physical health		**Clinical**	**Non-Clinical**
Increasing	TT	✓	
Decreasing	VST	✓	
	MT		✓
	OND		✓
	TX	✓	✓
	MDC	✓	✓
Predictor: SF-12 MCS Improved rating of mental health		**Clinical**	**Non-Clinical**
Increasing	TT	✓	
	VST	✓	
	MT		✓
Decreasing	TT		✓
	VST		✓
	MT	✓	

3.6 Exit Survey, Subjective Participant Responses

After completion of the task, participants were asked a series of questions regarding their experience. These included questions concerning their perception of their performance and workload during the task, their comfort with the equipment, and their opinions of the various interface manipulations.

Participants were positive about their experience. When asked to rate their comfort with using the handheld computer, the majority rated themselves as *very comfortable* (n=10), while two participants rated themselves as *comfortable,* and one individual rated them self as *neither comfortable nor uncomfortable*. None of the participants elected to report their experience as uncomfortable or very uncomfortable.

When asked to rate what they liked best about their experience, participants were in general positive about the task and technology, but negative with respect to the experimental equipment. Table 3.26 provides a sample of the participant likes and dislikes. The positive effect of using the playing card icons is observed in the 'likes' category, where participants related the task to playing a game, and made comments about their strategies, or how they liked the challenge. The aversions were, for the most part, related with their dislike for aspects of the experimental equipment, but nothing that could have interfered drastically with their performance. One exception was a participant who commented that they would have rather played a 'real game.' Furthermore participants were given breaks when necessary between the auditory and non-auditory feedback portions of the task to limit any neck strain or fatigue.

Table 3.26. Summary of participant likes and dislikes about their experience with the handheld and experimental task.

Likes	Dislikes
Nothing (n = 4)	*Nothing* (n = 4)
The ability to judge and distinguish cards	*It was an awkward position*
(I) appreciated that I got to ply a game	*Looking down (at the display) . . .I got stiff*
I liked the challenge of it	*I didn't like wearing the eye tracker . . .stupid thing*
To me, it was fun	*I didn't like when the sound was absent*
Moving the diamonds and hearts, because you can't miss those	*Not really my thing, but didn't dislike anything in particular*
	The glasses were slipping off my nose
	It went kind of fast . . . the pace was fast
	The very minute numbers and images
	I would have rather played a real game

In terms of their perceptions of their own performance and their interactions with the task, participants rated their overall performance, perceived difficulty of the task, their perception of how much effort was put forth to complete the task and their perception of the frustration they experienced. These questions provide an indication of participants' perceived mental workload, and were derived from subscales of the NASA TLX (NASA Ames Research Center 1987; Hart and Staveland 1988)(scales most easily understood by the participants in previous experiments).

Figure 3.18 (a-d) illustrates participants' responses to these questions of workload. Each workload factor was rated by the participants on a scale from 0 to 10, low to high. For performance, a score of 0 means the lowest, or worst possible perceived performance level, and 10 the absolute best performance. For Difficulty, 0 means not difficult at all and 10 means a maximum level of difficulty. For Effort, 0 signifies no effort and a value of 10 translates to the maximum amount of effort applied to complete the task. Finally, 0 means little to no frustration, and 10 means the maximum amount of frustration was experienced during the course of task completion.

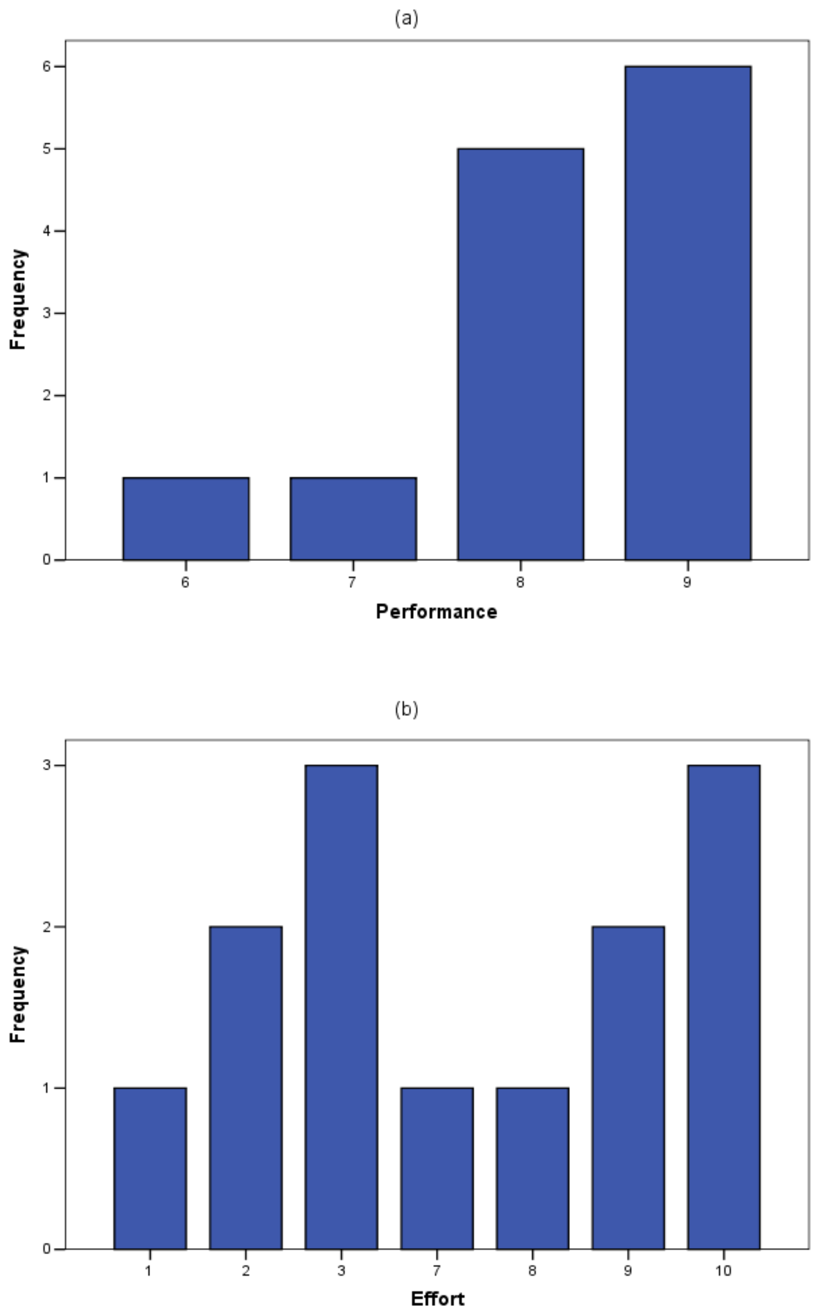


Figure 3.18. ((a)-(d)). Summary of responses to perceived workload subscales.

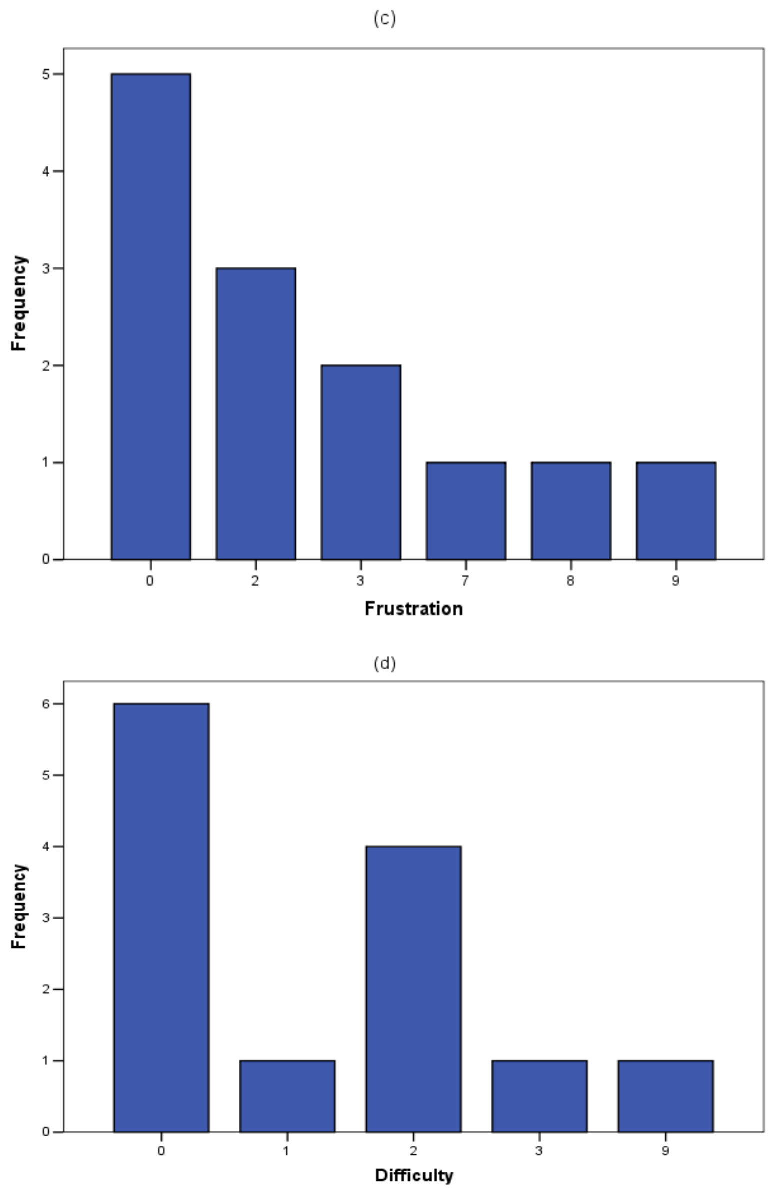


Figure 3.18. continued.

All of the participants rated his her overall performance on the task with a score of 6 or better, with the majority scoring their performance a 9 or 10. However, there was much more variability in the participants’ responses to perceived difficulty, effort and frustration. This is an accurate representation of the actual performance on the task, as measured by the accuracy and efficiency measures. Participants overall demonstrated a high success rate in task completion (correct card icon to correct to drop pile), but the rate at which they completed the different components of the task, and the occurrence of errors of commission during the task differed, dependent on personal and ocular factors.

As these are perceived rates of mental workload, it is of interest to compare these ratings to the measure of workload and information processing captured through eye movements and pupil diameter. To this end, the existing groupings for mean fixation duration, mean saccade duration, saccade to fixation ratio, and pupillary change were used to detect differences in the perceived task workload subscales for these groups.

None of these tests detected significant (alpha= .05) differences between these groups on any of the four subscales. In addition, non-parametric correlations were performed using Pearson's rho and Kendall's tau, neither of which detected significant difference between any of these groups for the perceived workload responses (alpha= .05). This is not surprising when considering the: (a) the small sample size and (b) low reliability in these perceived workload ratings when they are to divergent populations (apart from military or college students) (Moroney, Biers, Eggemeier and Mitchell 1992; Prinzel, Devries, Freeman and Mikulka 2001). That is, each participant's perception of performance, frustration, difficulty, etcetera, are subject to much internal bias. This provides evidence of the importance of the collection and analyses of the empirical and physiological date in the characterization of interactions.

The participants' reactions to the sound were consistent with its impact on performance. There was a mixture of reactions to it- both positive and negative. In expressing their comfort level with the sound, none of the participants rated their perception as *uncomfortable* or *very uncomfortable*. Participants only rated their experience with the sound as *very comfortable* (n = 5), comfortable (n = 7), and one rated their experience as *neither uncomfortable nor comfortable*.

The participants' response to question: *How helpful was the sound to your completion of the task?* resulted in even more varied response. Three of the participants rated the sound as *very helpful*, 2 rated the sound as *helpful*, 6 of the 13 responded that it was *neither helpful nor unhelpful*, 1 individual rated it as *unhelpful*, and the final participant rated the sound as *very unhelpful- or distracting*. The pie graphs in Figure 3.19 summarize these responses. In addition, the participants provided free responses on their general thoughts on the sound, which appear in Table 3.27, organized by positive, negative and mixed opinions.

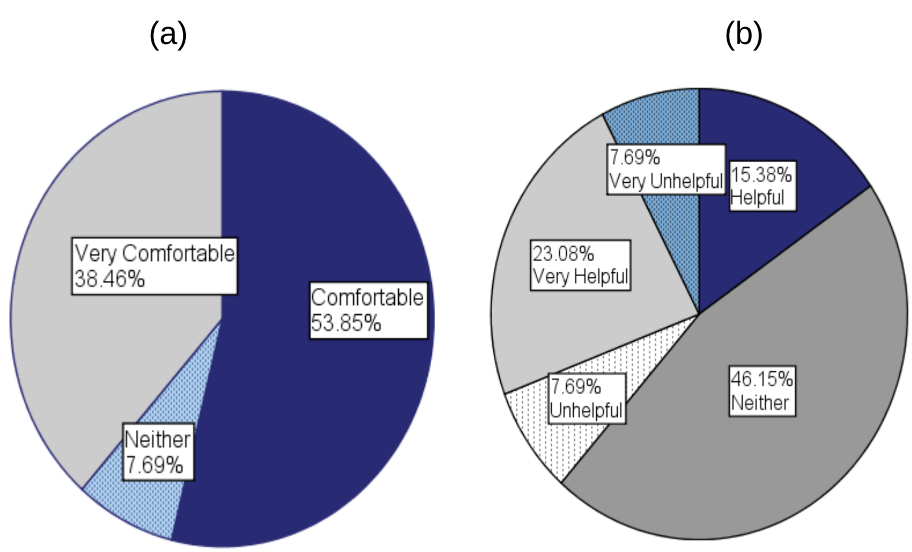


Figure 3.19. Summary of participant response to questions regarding (a) their comfort level with the auditory feedback and (b) their perception of the auditory feedback helpfulness.

Table 3.27. Participant opinions of auditory feedback, verbalized responses.

	Participant Opinions of the Auditory Feedback
Positive	*Like a trigger, like a teacher saying you're right.*
	Thought it was good-Very satisfying when you hear it.
	It sounded like my cat. . . I was very comfortable with it.
	It was fine.
	It was ok - fine
	It was ok - it didn't bother me
Negative	*Distracting*
	The light was spotted before the sound
	Surprised that it wasn't helpful; I didn't realize it wasn't helpful until it wasn't there
	It wasn't helpful for me - It would be more helpful for someone who has trouble seeing. It would also be helpful to associate the sound with the correct answers or have a specific sound for certain things.
	I was trying to avoid it - it was not really telling me I had achieved it (correct card to correct pile), and because it was making noise while I was trying to get somewhere - it needs to provide more information to be useful.
Neutral	*Tolerable*
	No opinion

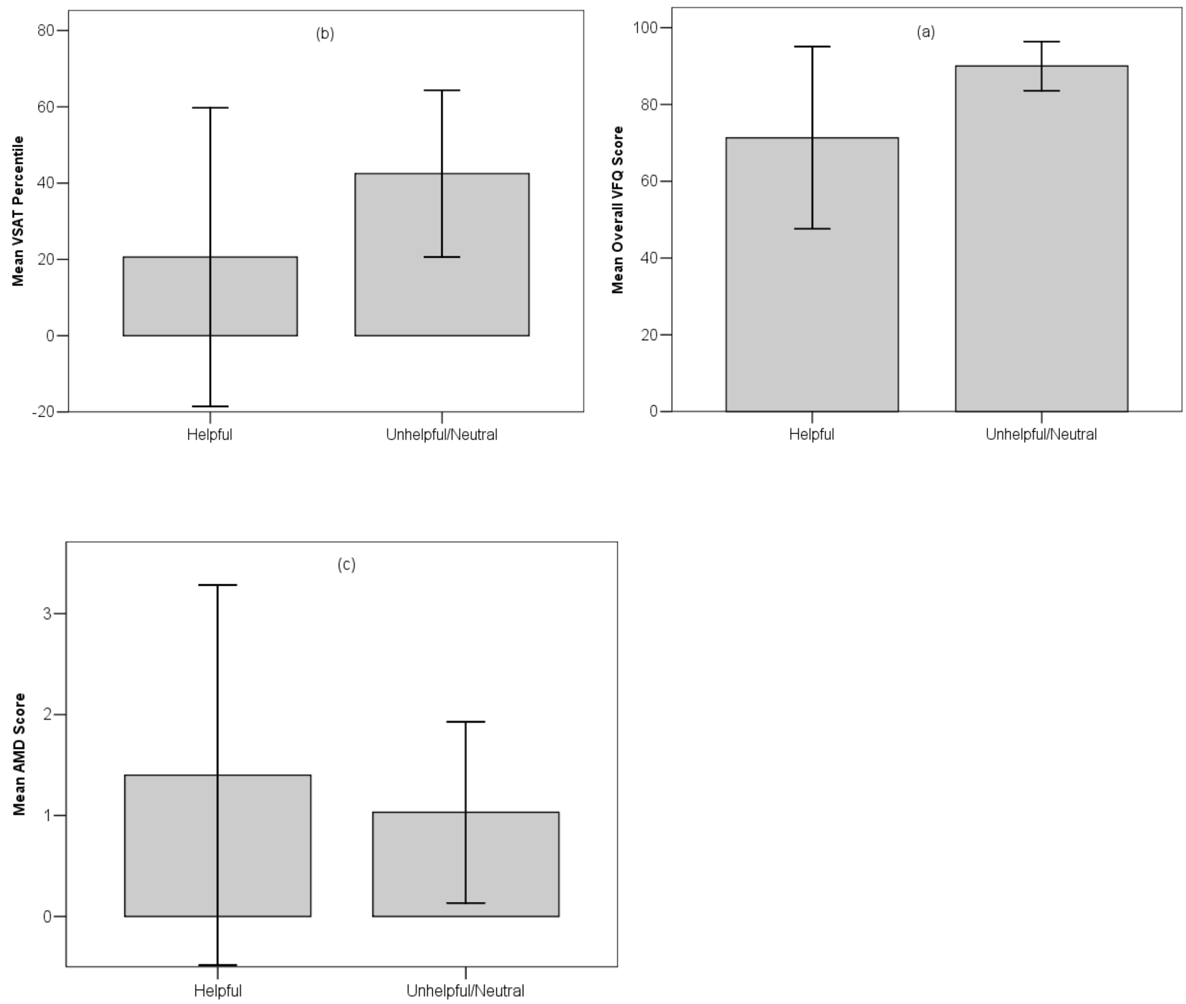


Figure 3.20. Comparisons of, VSAT, VFQ and AMD Score between participants based on their perception of auditory feedback helpfulness.

Further analyses were run on these perceptions of the auditory feedback, in light of the fact that the analyses of the performance metrics indicated that the auditory feedback provided the means to counteract the negative impact of disease and visual dysfunction experienced during the task. Analyses were applied to determine if those individuals who perceive the sound as helpful are actually those in the population who are likely to benefit from its presence. Based on their response to the helpfulness of the auditory feedback, participants were assigned to two groups. Group 1 consisted of those participants who responded as very helpful or helpful; Group 2 was comprised of participants who responded with *Neither*, *Unhelpful*, or *Very Unhelpful*.

Comparisons were made between these two groups using the Mann-Whitney test for non-parametric comparisons. Comparisons were made on VSAT, VFQ Overall, and AMD Score (summary values illustrated in Figure 3.20). Results revealed a significant difference only on VFQ overall between the two groups (Mann-Whitney U = 5.5, Z = -2.13, p = .034), and not on the other two measures of visual dysfunction (α = .05). In addition a Chi-Square test comparing controls and AMD with the two groups did not reveal significant differences bases on that factor. In terms of the VFQ, results demonstrated that individuals who rated the auditory feedback as helpful rated their perceived visual function and daily activities significantly lower, or worse than those individuals who were indifferent or felt the feedback was unhelpful. This suggests that individuals who rate their perceived visual function and daily activities as lower are more willing to consider the efficacy of non-traditional, non-visual supplemental cues to integrate with their residual vision.

3.7 Conclusions

The conclusions for the handheld portion of the experiment are discussed in terms of the relevant hypotheses established in Chapters 1 and 2. Tables 3.28-3.30 present the outcome related to each hypothesis, and the supporting evidence that emerged from the analyses. Each table contains a row with cells labeled *Efficiency, Subjective Response, Information Processing, and Subjective Response.* These refer to the class of measures that supported the conclusion. If a measure informed the conclusion, a check mark will appear in the appropriate column.

Table 3.28. Conclusions for Hypothesis 1, handheld experiment.

<u>Hypothesis 1</u>: For all users, there is a point of diminishing return for performance gains attributed to increases in set size and spacing.				
Contributing Results	**Efficiency**	**Accuracy**	**Information Processing**	**Subjective Response**
	✓	✓		
Supporting Evidence	*TT, VST, and the probability of AD errors all increased in the presence of increasing SS across all participants, for both clinical and non-clinical models. *DD increased in both clinical and non-clinical models in the presence of increasing ISp. *The probability of OND errors decreased with increasing ISp in both models. *The probability of TTHT increasing above the 85th %tile also decreased with increasing ISp, but was observed only in the clinical model.			
Conclusions	*While not surprising, the results confirms that there is a tradeoff between the amount of information presented on the screen at once, and increased visual search time, but to a degree that it influences the overall time to complete the task. *The presence of additional distracters imposed a higher probability for accuracy errors, also demonstrates the cost of increased information also interferes with the direct manipulation tasks, not only the visual search component. *While the impact of spacing on the distance traveled with the icon is intuitive, it did not impose an increase in the movement time. Spacing had a positive effect on the drop portion of the task. More accurate, efficient drops were facilitated through increased spacing between the drop folders.			

Table 3.28. continued.

<table>
<tr><td colspan="5">H1a. The point of diminishing return is dependent on a user’s visual capacity, including clinical measures of vision, visual attention, and subjective visual functioning.

H1b. Set size, and spacing will influence the components of the interaction (e.g., visual search, icon acquisition, dragging and dropping) in different ways and magnitudes, also dependent on the visual capacity of the participant.</td></tr>
<tr><td rowspan="2">Contributing Results</td><td>Efficiency</td><td>Accuracy</td><td>Information Processing</td><td>Subjective Response</td></tr>
<tr><td>✓</td><td>✓</td><td></td><td></td></tr>
<tr><td>Supporting Evidence</td><td colspan="4">*For participants with AMD, as the SS*AMD interaction term increased (increased disease severity, increased set size, or both), TT and VST were faster, and the likelihood for accidental drops decreased.

*Increases in the SS*VSAT and the ISp*VSAT interaction terms, like the SS*AMD, also affected TT and VST. As the score on the VSAT percentile increased, or measured visual attention improved, increases in SS and ISP imposed a less efficient, slower visual search termination and overall trial time.

*As SS*VFQ increased, the likelihood for TX decreased, while increases in ISp*VFQ decreased the likelihood of TTHT exceeding the 85th %tile. In other words improved perception of visual function improved the accuracy on the task.</td></tr>
<tr><td>Conclusions</td><td colspan="4">*While it would have been expected that the negative influence of SS and ISp would be more amplified with increases in visual function, these results indicate the opposite. However, it is justified in consideration of the nature of AMD.

*For those individuals with AMD, an increase in disease severity can lead to interruptions in visual field caused by blind spots, aberrations, or distortion. The number of icons they are able to see in their residual vision at one time may be limited. Therefore, increases to the number of icons on the screen may not be immediately perceived, until they reposition that useful field of vision. However, for those with full or improved visual function, it is more likely that they could perceive the entire display on the handheld at one time, at the onset of the task, and it was more likely to affect their visual search strategies.

*The results from SS*VFQ and ISP*VFQ demonstrate an improved accuracy for the participants with higher levels of perceived visual function especially in the presence of increased spacing or set size. However, individuals with AMD with increasing severity are not encumbered by increases in SS in terms of their efficiency in visual search and overall trail time. They are less apt in the precision of their approach to the drop pile (TX) and the benefits of spacing on the drop portion of the task are not as extreme in the release of the icon (TTHT) as for those visually healthy participants, the spacing and set size are more detrimental to icon manipulation and tracking with increasing levels of visual dysfunction.</td></tr>
</table>

Table 3.29. Conclusions for Hypothesis 2, handheld experiment.

<table>
<tr><td colspan="5"><u>Hypothesis 2</u>: The potentially positive influence of auditory feedback on the drag and drop task is effectively masked by the complexity of the task (multiple icons and multiple targets as compared to the single file – single folder task used in previous studies).</td></tr>
<tr><td rowspan="2">Contributing Results</td><td>Efficiency</td><td>Accuracy</td><td>Information Processing</td><td>Subjective Response</td></tr>
<tr><td>✓</td><td>✓</td><td></td><td>✓</td></tr>
<tr><td>Supporting Evidence</td><td colspan="4">The presence of AF in the non-clinical models generated main effects for faster MT, TT, and a decreased in likelihood for OND.

For the clinical models, there was a main effect for AF as it impacted a lower probability for TTHT to be in excess of the 85^{th} %tile.

The interaction term AFAMD, for the presence of AF and increasing disease severity, influenced measurable gains in performance - that is, decreased MT, DD, the probability for TTHT to exceed the 85th %tile, and decreased probability for OND, TX, and MDC errors.

The interaction between AFVFQ influenced increased TT, MT, and OND, suggesting that the impact of AF on those who perceive their vision as better, can be undesirable, and dampened the positive effects of AF observed across all participants, while not completely canceling it out.

The predictor term AFVSAT was influential on increasing the probability for AD, demonstrating a negative influence of AF on this portion of the task that was amplified in the presence of increasing VSAT Score.

The predictor term AFVSAT was influential on increasing VST, narrowing the performance gap that increasing VSAT introduce to the model of VT.</td></tr>
<tr><td>Conclusions</td><td colspan="4">*Several patterns emerged in the results concerning AF that confirm, but also enriched previous findings. Jacko and Colleagues (2004, 2005) demonstrated, for a simple drag and drop, that AF was the most helpful modality across participants, and furthermore did not cause any degradation of performance in any participants (including visually healthy).

*To the contrary, Vitense, Jacko and Emery (2003) observed detrimental affects of auditory feedback in a complex drag and drop task performance times were slower in the presence of the auditory cue. The participants in this study were exclusively visually healthy.</td></tr>
</table>

Table 3.29. continued.

Conclusions	*This present study on the handheld demonstrated both potential decrements to performance, but large performance gains based on the ocular pathology of the participants. Showing that AF is of use with multiple targets, but should not be universally incorporated across all users, depending on their visual profiles. While in some cases it can be helpful to all participants, ability of AF to universally help all participants was not a clear cut result in this study, and in some cases showed signs it could be detrimental to performance in OND errors, and VS. AF and other supplemental non-visual cues should be judiciously incorporated into an interface, based on the participants' visual profile.

Table 3.29. continued.

H2a. The presence of auditory feedback will not affect the visual search component of this task				
Contributing Results	**Efficiency**	**Accuracy**	**Information Processing**	**Subjective Response**
	✓	✓		✓
Supporting Evidence	*The AF*VFQ interaction term emerged as an influential factor in increasing VST. *Participants with higher VFQ scores were less likely to rate the feedback as helpful to their task.			
Conclusions	*The use of supplemental non-visual cues can be distracting to those individuals with lower or no visual dysfunction while using a small interactive display, to the point where it interferes with task components separate from those components associated with the sound.			

H2b. As the number of icons and potential drop targets increase, the presence of auditory feedback can have detrimental effects on the interaction.				
Contributing Results	**Efficiency**	**Accuracy**	**Information Processing**	**Subjective Response**
	✓	✓		
Supporting Evidence	*Increases in SS affected only TT, VST and AD - the search and initial approach to the drop pile. *VST and the likelihood for ADT were also shown to increase in the presence of improving vision in condition where the interaction term VSAT*AF increased.			
Conclusions	*When combined, the increase of SS and presence of AF can cause decrements to interaction accuracy and efficiency, but only for those participants with lower levels of dysfunction. Again this leads to the conclusion that the decision to include auditory feedback must be made based on an individual's specific, unique visual profile.			

Table 3.30. Conclusions for Hypothesis 3, handheld experiment.

Hypothesis 3: Measures of ocular health and visual function are predictors of performance in the required task.				
Contributing Results	**Efficiency**	**Accuracy**	**Information Processing**	**Subjective Response**
	✓	✓	✓	✓
Supporting Evidence	*Measures of ocular health were the most influential predictors of performance in several instances, observed through several main effects of the clinical and non-clinical ocular factors, as well as via the interactions with the interface manipulations. *Both AMD Score and CS were included predictor in all efficiency models, and accuracy models. *NVA was an included predictor for TT, MT, DD, TTHT, and all efficiency measures. *VFQ was an included predictor of DD, MDC, AD, and TX. *VSAT was an included predictor for TT, VST and TT. *Differences in mean fixation duration were linked with lower CS. *The perception of feedback helpfulness was associated with participants' decreasing perception of their visual function and daily activities (VFQ).			
Conclusions	*The magnitude of the impact of the visual factors on the models, as well as their persistence in the models demonstrates the critical impact of visual function in the manipulation of GUI's of any size, even handhelds. *In the majority of instances, degradation to the visual capacity, or the self perceived notion of degradation imposed quantifiable impositions on effective interaction with the handheld.			

Table 3.30. continued.

H3a. Certain components of the interaction (e.g., the visual search, drag and drop) are more susceptible to the negative impacts of limited ocular functioning. **H3b.** The components of the interaction are influenced by each measure of ocular health and visual function to a different degree.				
Contributing Results	**Efficiency**	**Accuracy**	**Information Processing**	**Subjective Response**
	✓	✓	✓	
Supporting Evidence	*Clinical, non-clinical predictors, or their combined interaction effect were the most influential across all models with few exceptions; at least one clinical factor affected each model. *There were no main effects of any visual predictors (only interactions) for the non-clinical models of OND, or TTHT. *For both clinical and non-clinical models of DD, the layout of the interface had a more influential effect on the distance traveled with the icon. The row, and spacing were the most influential, and this was followed by dexterity in the clinical model, and DL in the non-clinical model.			
Conclusions	*The emerging trend from this data is that visual factors are the most influential on the visual search component of the task, and in the case of the clinical models, on the precision executed in the drop portion of the task. *The movement and visual tracking of the icon across the display, using the stylus is more prone significantly impacted by issues of screen layout and personal factors such as physical health and dexterity. *However, the global impact of the visual factors is emphasized in the model of TT, which accounted for VFQ, VSAT, AMD score, and interaction terms. That said, in designing interfaces for individuals with visual impairments priority should be given to components of the task that are characterized by visual search, or targeting. These components of the task have the most potential for measurable gains, independent of extraneous factors.			

Table 3.30. continued

<table>
<tr><td colspan="5">H3c. Different measures of ocular ability can delineate which components of the task will be executed in a less efficient manner, following the speed-accuracy tradeoff common to most HCI tasks.</td></tr>
<tr><th rowspan="2">Contributing Results</th><th>Efficiency</th><th>Accuracy</th><th>Information Processing</th><th>Subjective Response</th></tr>
<tr><td>✓</td><td>✓</td><td>✓</td><td></td></tr>
<tr><th>Supporting Evidence</th><td colspan="4">*Several instances of the speed accuracy trade off emerged in the analyses of the handheld data for both the efficiency and accuracy measures as well as both the visual search and icon manipulation portions of the task, and these were associated with ocular factors.

*As VFQ increased, improving the perception of visual function the probability of MDC, DD, and AD decreased but the likelihood for TX increased. While not a speed accuracy trade off, per se, there appears to be a tradeoff in terms of the path efficiency (the most direct path to the drop pile would incur no TX); while the participant was very accurate in their approach, not changing directions or releasing the icon prematurely.

*As NVA declined in the models, MT was slower, but DD was shorter. Participants with more difficulty focusing on near vision were more precise in the approach to the drop pile, but at the expense of delayed movement time.

*As CS improved, TT and VST and MT were all faster, while DD and the probability for AD and MDC all increased. Participants with better CS were moving faster but at the price of more errors on the approach to the drop pile.

*The slower rates of VST, TT, and MT for degraded CS scores were accounted for by the analyses of fixation duration and pupillary response. Individuals with significantly poorer CS also had a significantly greater fixation duration and pupillary response over the course of the task.</td></tr>
<tr><th>Conclusions</th><td colspan="4">*The effects of increasing AMD severity impose degradation in performance on all phases of the task.

*The impact of the more specific ocular measures, CS and NVA, however are not completely straightforward, as tradeoffs emerged in the outcome variables as either factor deteriorated. The priorities of the specific task (accuracy vs. speed) must be considered in the development, especially when deciding if and how to accommodate the needs of individuals with diminished CS and NVA. Those tasks with the contextual constraints time/efficiency would undoubtedly be susceptible to erroneous interactions for this population.

*It is important to consider that the display for the experimental task was designed promote a high level of contrast sensitivity, based on the results of several previous studies touting the importance of CS in display design for the visually impaired. The effects of CS are undeniably amplified on more typical, low contrast displays.</td></tr>
</table>

Table 30. continued.

H3d. The predictive power of attributes of visual function used to classify the outcome of interactions will differ in terms of the amount of influence on the task.				
Contributing Results	**Efficiency**	**Accuracy**	**Information Processing**	**Subjective Response**
	✓	✓	✓	
Supporting Evidence	*The clinical visual factors strongly influenced all components of the task, except for DD, and saccade based metrics. The specific clinical factor with the most influence change between models, but all were highly indicative of performance *NVA was an extremely influential factor on TX, OND, and MDC, and the second most influential factor in decreasing the likelihood for AD. *VSAT was more often a predictor in the models of TT, VST, MT, but not TTHT, DD, or any of the accuracy measures. VSAT was the most influential factor in both VST and TT. *VFQ was more commonly included as a main effect in the models of error likelihood, MDC, TX and AD, and was the most influential factor in AD. *CS was the only significant in the increased levels of fixation duration. *VFQ was the only significant predictor of participants' perception of the helpfulness of the feedback.			
Conclusions	*NVA and CS directed the accuracy of the task and the movement of the icon to a greater degree than AMD Score (in all cases but AD). *This suggests that the negative effects of AMD have more bearing on the identification of the target icon amongst distracters, while CS and NVA should definitely be considered in the movement and targeting of icons (while this relationship is not exclusive). *Future studies should consider the interactions of all three visual factors and their interaction with the interface components. *The VSAT is a powerful assessment of visual search efficacy, while the VFQ can indicate the ability for the person to integrate supplemental non-visual cues for improved interaction in performance.			

Table 30. continued.

<table>
<tr><td colspan="5">H3e. The influence of ocular functioning on task interaction greatly overrides other normal, age-related declines in mental and physical health.</td></tr>
<tr><td rowspan="2">Contributing Results</td><td>Efficiency</td><td>Accuracy</td><td>Information Processing</td><td>Subjective Response</td></tr>
<tr><td>✓</td><td>✓</td><td>✓</td><td>✓</td></tr>
<tr><td>Supporting Evidence</td><td colspan="4">*The relative magnitude of the coefficients for the visual factors in the clinical models were consistently the most influential on the outcome variables with the exception of the Model for DD, where the user's dexterity weighed more heavily than NVA, CS, or AMD.

*While AMD Score was the overriding predictor in the TT, and VST for the clinical models, dexterity was slightly more influential on TT than NVA, and SF-12 PCS, MCS and Dexterity were more influential on VST than CS.

*The non-clinical model for TT, SF-12 mental health, age and dexterity were more influential on TT than VFQ.

*While ocular health (CS) was related to longer mean fixation duration, while SF-12 was indicative of less efficient search strategies, modeled through the saccade measures.

*Personal factors were not accountable for any of the subjective opinions about the task, while VFQ was related the participants' opinion of the helpfulness of the auditory cue.

*Both VFQ and SF-12 PCS drove the non-clinical model for MDC equally; visual factors did not dominate this model.</td></tr>
<tr><td>Conclusions</td><td colspan="4">*The non-visual factors, while they are significant predictors included in the model, had opposing influence between the clinical and non-clinical models. This suggests that the different classes of ocular factors, particularly the clinical factors, likely account for some of the expected age-related and general health related performance differential, in a way that the non-clinical factors did not.

*The variance associate with age, dexterity, and overall mental and physical health should be accounted for in predicting interaction ability, but visual ability is presents a more critical influence beyond the normally anticipated inter-user differences.

*Future studies need to consider the interactions between visual factors and these personal factors to fine tune how the interface should accommodate the individual's needs.</td></tr>
</table>

Table 30. continued.

H3f. Non-clinically-acquired measures of visual function, such as the users' perception of the impact of visual dysfunction on their activities of daily living (VFQ), and functional visual attention are more powerful predictors of the outcomes of the interaction than clinical factors.				
Contributing Results	**Efficiency**	**Accuracy**	**Information Processing**	**Subjective Response**
	✓	✓	✓	✓
Supporting Evidence	*The non-clinical factors, VFQ, and VSAT were more difficult to interpret in the models, and were not consistently influential on the models as the clinically acquired factors. * The clinical models of TTHT, and the accuracy measures were not as rich as the clinical models; the variance in exp(B) was very slight between the factors, making it difficult to draw definitive conclusions on these points (although they were highly consistent with the clinical models). *Non-clinical models of efficiency, particularly for VST, TT and MT were highly informative in characterizing the interaction, incorporating both main effects and interactions of the visual factors, interface characteristics. *The R-squared values for the linear models were consistently higher for the clinically acquired- factors, than the non-clinically-acquired. In addition the HL value for the non-clinical MDC model was also significant, indicating a poor fitting model. The percentage of correctly classified cases was higher for the clinically-acquired models.			

Table 3.30. continued.

Conclusions	*The non-clinical assessments of vision are informative, but not wholly indicative of the interaction, as demonstrated with the handheld results. The VSAT held more predictive value in the assessment of visual search, and VFQ in the efficacy of AF on the models, and the accuracy measures. *The VFQ appears to be indicative of users' acceptance of non-traditional assistive technology and also potentially their ability to incorporate it into their interactions to experience performance gains that counteract the effects of visual dysfunction. *The use of non-clinical assessments of visual health have potential in the eventual ability to assess the participants' interactions and make changes to the computer interface in the absence of current (more expensive) clinically acquired metrics. *There is room for further development and understanding of the relationships of the non-clinical measures with both computer interaction and the clinically-based factors. *The clinically acquired measures presented a more comprehensive portrayal of their impact on the various components of the interaction, and their construct, are more easily understood in terms of the quantifiable impact.

3.8 References

Backs, R. W. and L. C. Walrath (1992). "Eye movement and pupillary response indices of mental workload during visual search of symbolic displays." Applied Ergonomics **23**: 243-254.

Bird, A. C., N. M. Bressler, S. B. Bressler, I. H. Chisolm, G. Coscas, M. D. Davis, P. T. V. M. d. Jong, C. C. W. Klaver, B. E. K. Klein, R. Klein, P. Mitchell, J. P. Sarks, S. H. Sarks, G. Soubrane, H. R. Taylor and J. R. Vingerling (1995). "An international classification and grading system for age-related maculopathy and age-related macular degeneration: The international ARM epidemiological study group." Survey of Ophthalmology **39**(5): 367-374.

Darling, W. G., J. D. Cooke and S. H. Brown (1989). "Control of simple arm movements in elderly humans." Neurobiology of Aging **10**: 149-157.

Edwards, P. J., L. Barnard, V. K. Emery, J. S. Yi, K. P. Moloney, T. Kongnakorn, J. A. Jacko and F. Sainfort (2005). "Understanding users with Diabetic Retinopathy: Factors that affect performance in a menu selection task." Behaviour & Information Technology **24**(3): 175-186.

Edwards, P. J., L. Barnard, V. K. Emery, J. S. Yi, K. P. Moloney, T. Kongnakorn, J. A. Jacko, F. Sainfort, P. R. Oliver, J. Pizzimenti, A. Bade, G. Fecho and J. Shallo-Hoffmann (2004). Strategic Design for Users with Diabetic Retinopathy:
Factors Influencing Performance in a Menu-Selection Task. Sixth International ACM SIGCAPH Conference on Assistive Technologies (ASSETS 2004), Atlanta.

Fahrmeir, L. and G. Tutz (1994). Multivariate Statistical Modelling Based on Generalized Linear Models. New York, Springer-Verlag.

Field, A. (2000). Discovering Statistics Using SPSS for Windows. London, Sage Publications.

Fitts, P. M. (1954). "The information capacity of the human motor system in controlling the amplitude of movement." Journal of Experimental Psychology **47**: 381-391.

Hart, S. G. and L. E. Staveland (1988). Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. Human Mental Workload. N. Meshkati. Amsterdam, The Netherlands, Elsevier Science Publishers B. V. (North-Holland)**:** 139-183.

Jacko, J. A. (2000). "Visual impairment: The use of visual profiles in evaluations of icon use in computer-based tasks." International Journal of Human-Computer Interaction **12**(1): 151-164.

Jacko, J. A., L. Barnard, T. Kongnakorn, K. P. Moloney, P. J. Edwards, V. K. Emery and F. Sainfort (2004). "Isolating the effects of visual impairment: Exploring the effect of AMD on the utility of multimodal feedback." CHI Letters **6**(1): 311-318.

Jacko, J. A., A. B. Barreto, I. U. Scott, J. Y. M. Chu, H. S. Vitense, F. T. Conway and W. B. Fain (2002). "Macular degeneration and visual icon use: Deriving guidelines for improved access." Universal Access and the Information Society **1**: 197-296.

Jacko, J. A., V. K. Emery, P. J. Edwards, M. Ashok, L. Barnard, T. Kongnakorn, K. P. Moloney, F. Sainfort and I. U. Scott (2004). "The effects of multimodal feedback on older adults' task performance given varying levels of computer experience." Behaviour & Information Technology **23**(4): 247-264.

Jacko, J. A., K. P. Moloney, T. Kongnakorn, L. Barnard, P. J. Edwards, V. K. Emery, F. Sainfort and I. U. Scott (2005). "Multimodal feedback as a solution to ocular disease-based user performance decrements in the absence of functional visual loss." International Journal of Human-Computer Interaction. **18**(1).

Jacko, J. A., K. P. Moloney, T. Kongnakorn, L. Barnard, P. J. Edwards, V. K. Leonard, F. Sainfort and I. U. Scott (2005). "Multimodal feedback as a solution to ocular disease-based user performance decrements in the absence of functional visual loss." International Journal of Human-Computer Interaction **18**(1): 183-218.

Jacko, J. A., R. H. Rosa, I. U. Scott, C. J. Pappas and M. A. Dixon (1999). Linking visual capabilities of partially sighted computer users to psychomotor task performance. 9th International Conference on Human-Computer Interaction, Munich, Germany.

Jacko, J. A., I. U. Scott, F. Sainfort, L. Barnard, P. J. Edwards, V. K. Emery, T. Kongnakorn, K. P. Moloney and B. Zorich (2003). "Older adults and visual impairment: What do exposure times and accuracy tell us about performance gains associated with multimodal feedback?" CHI Letters **5**(1): 33-40.

MacKenzie, I. S., T. Kauppinen and M. Silfverberg (2001). "Accuracy measures for evaluating computer pointing devices." CHI Letters **v3**: 9-16.

Miller, S. M. (1987). "Monitoring and blunting: validation of a questionnaire to assess styles of information seeking under threat." Journal of Personally and Social Psychology **52**(2): 345-353.

Moroney, W. F., D. W. Biers, F. T. Eggemeier and F. A. Mitchell (1992). A comparison of two scoring procedures with the NASA task load index in a simulated flight task. IEEE 1992 National Aerospace and Electronics Conference (NAECON 1992), Dayton, OH, IEEE.

NASA Ames Research Center (1987). NASA Human Performance Research Group Task Load Index (NASA-TLX) instruction manual. Nasa Ames Research Center.

Neter, J., M. H. Kutner, C. J. Nachtsheim and W. Wasserman (1996). Applied Linear Statistical Models. Boston, McGraw-Hill.

Pallant, J. (2003). SPSS Survival Manual. Philadelphia, Open University Press.

Pew, R. W. (1969). "The speed-accuracy operating characteristic." Acta Psychologica **30**: 16-26.

Prinzel, L. J. I., H. Devries, F. G. Freeman and P. Mikulka (2001). Examination of automation-induced complacency and individual difference variates. Hanover, MD, NASA**:** 43.

Scott, I. U., W. J. Feuer and J. A. Jacko (2002). "Impact of graphic user interface screen features on computer task accuracy and speed in a cohort of patients with Age-related Macular Degeneration." American Journal of Ophthalmology **134**(6): 857-862.

Scott, I. U., W. J. Feuer and J. A. Jacko (2002). "Impact of visual function on computer task accuracy and reaction time in a cohort of patients with Age-related Macular Degeneration." American Journal of Ophthalmology **133**(3): 850-857.

Vitense, H. S., J. A. Jacko and V. K. Emery (2002). "Foundation for improved interaction by individuals with visual impairments through multimodal feedback." Universal Access and the Information Society **2002**(2): 76-87.

Vitense, H. S., J. A. Jacko and V. K. Emery (2003). "Multimodal feedback: an assessment of performance and mental workload." Ergonomics **46**(1-3): 68-87.

CHAPTER 4

DESKTOP PC:
ANALYSES, RESULTS & CONCLUSIONS

4.1 Introduction

This chapter presents the analyses, results and conclusions associated with the different facets of the icon search, selection and manipulation task executed on a desktop computer with a mouse input device. Fourteen volunteers from the NSU College of Optometry patient pool and associates of NSU staff participated in the desktop component (distinct from those who participated in the handheld component). Twelve participants' only ocular diagnosis was some level of AMD, and the other two were age-matched controls, not diagnosed with any ocular disease. The demographic backgrounds and ocular profiles of these individuals were summarized in Chapter 2, together with the experimental procedures and variables considered in the desktop PC task.

Overall, this set of participants' success with the desktop PC task was not as prevalent as the participants' interactions on the handheld. Whereas all the participants but one completed the 162 trials on the handheld PC, three of the participants in the desktop PC task skipped trials because they articulated an inability to work with the smaller icons sizes and/or in some cases communicated fatigue with the overall task. Even so, eleven of the participants completed all 162 trials, while the other three participants completed 21, 54, and 87 trials respectively.

Amongst these participants, 324 trials were skipped, in total. These skipped trials were evenly distributed across the three spacing conditions (ISp); the three set size conditions (SS) and the two auditory feedback conditions. There was, however, a disproportionate number of trials skipped which contained the smallest icon, as illustrated in Figure 4.1. This effect was not observed for the handheld condition, because icon size was not manipulated. The data from the participant who completed only 21 of the 162 trials was excluded from analyses and interpretation (This individual was diagnosed with AMD). These data were excluded as outlying cases because all of trials for this individual were extraordinarily long compared to the other participants (on the order of 4 minutes per drag and drop). In addition, despite the fact that participants were screened for computer experience, this subject's performance, verbalizations, and comfort level in the task led the experimenters to suspect otherwise.

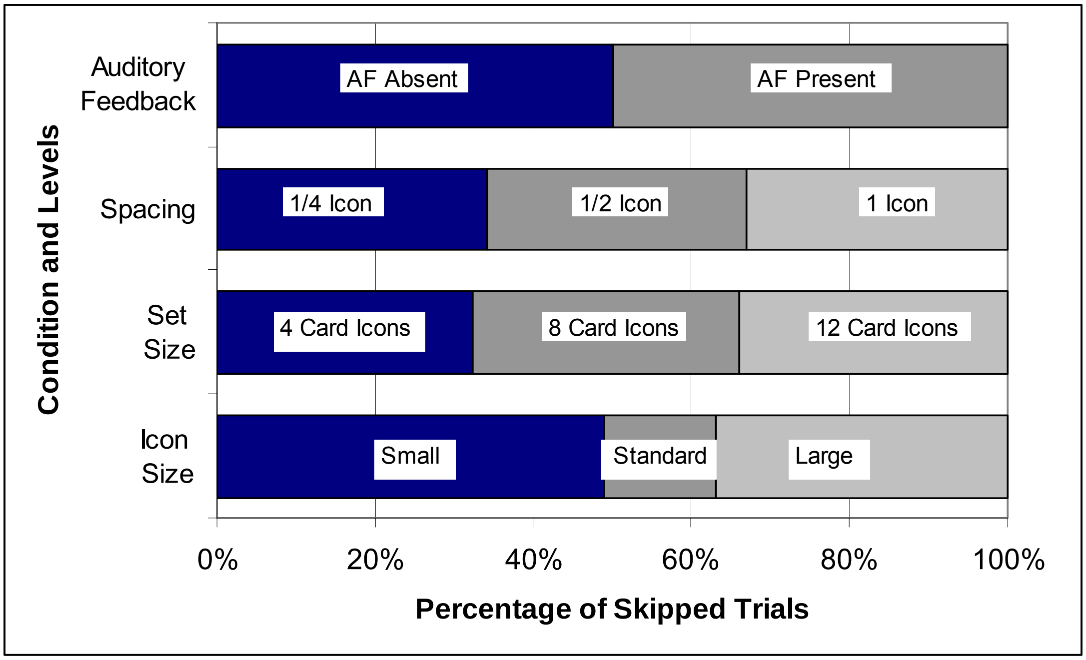


Figure 4.1. Breakdown of skipped trials.

Overall, the participants executed the task with a high degree of success. Across all participants, 95.6% of the trials were correctly completed. Participants worked in the desktop PC condition on average 15.31 minutes (standard deviation = 5.63 min; median = 14.75). The high success rate implies that global summary performance measures for this task are not highly informative in the explanation of the various ocular, interface, and participant-related factors' influence on the interaction. For this reason, the measures of the various phases of implicit and explicit interaction were of particular interest in providing a more illustrative account of task interaction.

The analyses for the desktop PC task are, in general, consistent with those applied to handheld in order to facilitate comparisons between the two platforms. Unique to the desktop analysis was the consideration of errors in the acquisition of the card icon (IA). Missed opportunities were analyzed for both the selection of the card icon, and the release of the icon into the drop pile. The desktop component also controlled for an additional independent task-related variable: icon size. Analyses were run on the *Efficiency*, *Accuracy*, and *Information Processing* outcome measures using the same set of statistical tools as was applied to the handheld: linear regression, logistic regression, and non-parametric comparisons. The variables considered in modeling the interaction metrics were consistent, with the inclusion of an additional task-related factor, icon size (Sz, 3 levels).

The analyses and results presented in this chapter are divided into three sections according to the three classes of dependent measures considered: 1) Efficiency, 2) Accuracy, and 3) Information Processing. Each dependent measure is considered in light of clinically-acquired visual factors, non-clinically-acquired visual factors, personal characteristics, interface features (i.e., icon size, set size, spacing, and auditory feedback) and other extraneous task-related factors (e.g., location of target icons, and order effects). Finally, the qualitative data gathered from the participants in exit surveys are summarized and related back to the more empirically-driven results. This chapter concludes with a section linking the results back to those hypotheses relevant to the desktop PC task.

Tables 4.1-4.4 summarize the predictor variables considered in modeling the variance of the dependent variables, grouped according to *task-related factors, general participant-related factors*, *clinically-acquired ocular factors,* and *non-clinically-acquired ocular factors*. The VFQ measures are further summarized by Figure 4.4, which details the responses on the twelve subscales and overall score.

Because of the lack of diagnostic specificity in the individual VFQ subscales for this population sampled, only the VFQ Overall Score was employed in the statistical analyses (the inclusion of all subscales which are so highly correlated would have otherwise compromised the robustness of the analyses). Also included in these summary tables is a description of the levels observed for each predictor in the desktop participant group. In addition to those variables listed, statistical interactions were considered for AMD Score, VSAT and VFQ Overall Score with the controlled interface factors (SS, AF, Sz and ISp).

Table 4.1. General participant-related factors.

Predictor Variable	Description	Observed Levels
Age	Participant's age at the time of the study	Range: 62-90 years Mean: 76.5 Median: 77
SF-12 Physical Component Score *SF-12 PCS*	Survey of participant's self reported physical health at the time of the experiment, from 1 (worst) to 100 (best)	Range: 20.13-60.56 Mean : 43.93 Median : 45.31
SF-12 Mental Component Score *SF-12 MCS*	Survey of participant's self-reported mental health at the time of the experiment, from 1 (worst) to 100 (best)	Range: 28.37-63.90 Mean: 44.93 Mean: 44.89
Purdue Pegboard Test of Manual Dexterity *Dexterity*	A count of the average number of pins inserted into small holes in a board over 3, 30 second trials, from 0 (worst) to 30 (best)	Range: 2-13 Mean: 8.88 Median: 9.5

Table 4.2. Clinically-acquired ocular factors.

Predictor Variable	Description	Observed Levels
LogMar Near Visual Acuity† *NVA*	Ability to visually focus on fine details at a distance of 40 cm, translated from Snellen acuity (e.g. 20/20). Scores can range from .1 (best) to 1(worst)	Range:0.00-1.00 Mean: .401 Median: .238
Contrast Sensitivity† CS	Measure of how visible an image is before it becomes indistinguishable from a uniform field, from 0 (low) to 60 (high)	Range: 7.5-36.0 Mean: 27.38 Median: 30.63
AMD Severity Score† *AMD Score*	A diagnosis of severity of disease based on examination of photographs of the eye from no disease (0) to severe (4)	Range: 0-4.00 Mean: 1.89 Median: 1.000

†For NVA, CS and AMD Severity Score, weighted average of the best and worst eye (.75 * best + .25 * worst) approximated binocular visual field.

Table 4.3. Non-clinically-acquired ocular factors.

Predictor Variable	Description	Observed Levels
Visual Attention Test *VSAT*	A paper based assessment of sustained visual attention. Scores on the VSAT have been compared to an age-related normative sample, and recoded according to their relative percentile (based on age). The higher the percentile, the less severe the detected impairment. -At or below 2nd percentiles: Significant, impaired vision -At or between 3rd to 16th percentile: Suggestive, borderline impairment -At or above 17th percentile: Within the normal range of performance	Range: 1st -58th percentiles Mean: 20.00 Median: 18.00 <u>Incidence of categorical classification:</u> Significant: 4 Borderline: 3 Normal: 7
NEI Visual Functioning Questionnaire-25 *VFQ*	Self-perceived assessment of visual function and daily activity, based on responses to the verbally administered NEI-VFQ-25. Scores are generated on 13 subscales, and 1 overall score and range from 0 (maximum interference with daily functioning, or wore perceived visual function) to 100 (no interference, best possible perceived visual function).	Overall VFQ Score Range: 19.04-98.08 Mean: 71.43 Median: 78.27 (95%CI are provided for Overall Score and the other 13 subscales in Figure 4.2)

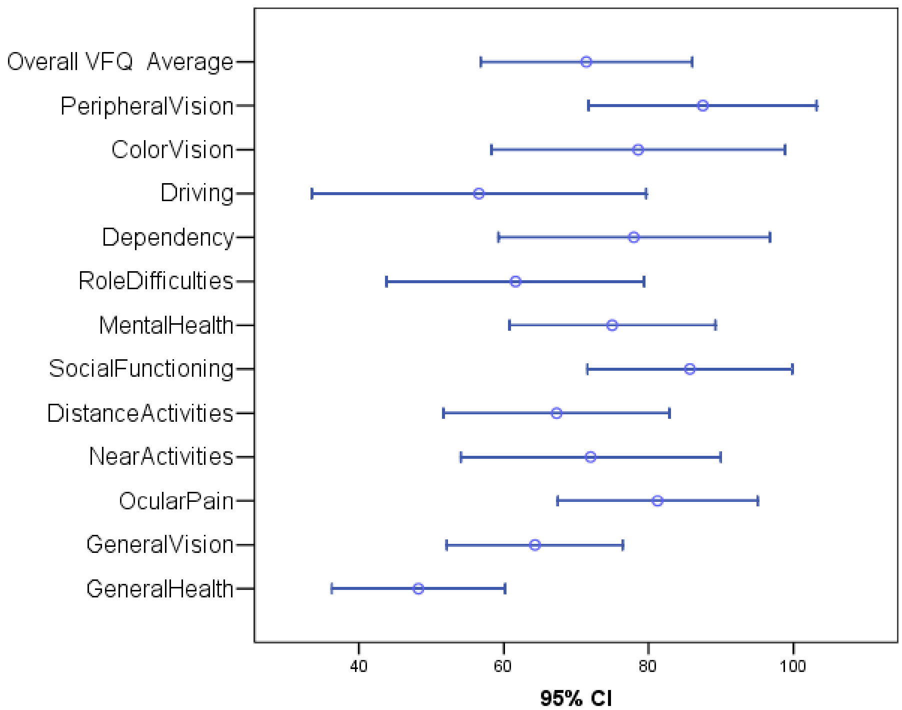


Figure 4.2. 95% Confidence intervals for the 13 VFQ subscales and overall average.

Table 4.4. Task-related factors.

Predictor Variable	Description	Observed Levels
Set Size (SSI)	The number of card icons presented for each trial	1: 4 card icons 2: 8 icons 3: 12 icons
Inter-Icon Spacing (ISp)	The 'white' space between the icons	1: ¼ icon 2: ½ icon 3: 1 icon
Icon Size (Sz)	The size of the card icons and drop pile	1: Small icons (handheld size) 2: Standard windows size 3: Large icons (Recommend for users with visual impairments)
Auditory Feedback (AF)	Supplemental auditory feedback to communicate the position of the card for an accurate drop	0: AF absent 1: AF present
Column Location (Column)	The column where the target card icon is located for each trial	1: leftmost 2: middle 3: rightmost
Row Location (Row)	The row where the target card icon is located for each trial	1: top 2: 2nd from top 3: 2nd from bottom 4: bottom
Drop Location	The row number of where the correct drop pile for each trial was located	1: top 2: 2nd from top 3: 2nd from bottom 4: bottom
Trial Number (Trial #)	Sequential position of the trial within a participant's overall experimental session	Range: 0 -161

In the following sections, the specific analyses executed and the results are detailed, anchored in the taxonomy of dependent variables: *Efficiency, Accuracy, and Information Processing*. Each set of dependent measures is modeled using clinically-acquired ocular factors and non-clinically-acquired ocular factors separately, while each model consistently considers general participant-related factors (e.g., age) and task-related factors (e.g., spacing, set size, trial number). The predictor variables considered in the clinically-acquired ocular analyses were consistent between regressions on each outcome measure, as were those considered between the non-clinically-acquired prediction models. These were consistent so that the relative effects between different task phases could be compared. A complete aggregated list of predictor variables follows:

Ocular Factors	Personal Factors	Interface Factors	Interaction Terms
Clinical	SF-12 PCS	SS	AMD Score * SS
NVA	SF-12 MCS	Sz	AMD Score * Sz
CS	Dexterity	ISp	AMD Score * ISp
AMD Score	Age	AF	AMD Score * AF
Non-Clinical		Column Location	VFQ * SS
VSAT Percentile		Row Location	VFQ * SZ
Overall VFQ Score		Drop Location	VFQ * ISp
		Trial Number	VFQ * AF
			VSAT * SS
			VSAT * Sz
			VSAT * ISp
			VSAT * AF

4.2 <u>Efficiency Measures</u>

General Summary

A complete inventory of the efficiency outcome measures collected for the desktop and the distributions are provided in Table 4.5. Figures 4.3 and 4.4 summarize the nature of the task through graphs of the means and standard error for each of the efficiency measures. Figure 4.3 illustrates that relative to TT, the majority of participants' time in each trial was spent in VS. The relative difference however, between VS and MT was not as large as what was observed in the handheld participant interaction. While TTHT and FTHT, as shown in Figure 4.3, are much smaller in scale than VST and MT, the two measures are proportionally greater than the highlight measures observed in the handheld task. Figure 4.4 presents a summary of the efficiency measures, which are based on the distance traversed to move the card icon to the drop pile, MV and ME.

Table 4.5. Summary of efficiency measures of desktop interaction and abbreviation.

Efficiency Measure	
Trial Time-*TT* (msec)	Mean (Std. Error): 6058.60 (131.45) Median: 4486.5 Minimum:1311 Maximum: 80486
Visual search time-*VST* (msec)	Mean (Std. Error): 2324.53 (54.47) Median: 1682 Minimum: 210 Maximum: 37153
Movement Time-*MT* (msec)	Mean (Std. Error):1829.81 (21.87) Median: 1713 Minimum: 50 Maximum: 9324
Final target highlight time-*FTHT* (msec)	Mean (Std. Error): 1001.74 (1.77) Median: 831 Minimum: 0 Maximum: 22793
Final target highlight time-T*THT* (msec)	Mean (Std. Error): 1169.34 (29.18) Median: 942 Minimum: 20 Maximum: 25267
Movement Variability-*MV* (pixels)	Mean (Std. Error): 46.63 (1.53) Median: 20.13 Minimum: <.001 Maximum: 502.933
Movement Error-*ME* (pixels)	Mean (Std. Error): 20.11 (.680) Median: 8.13 Minimum: <.001 Maximum: 230.467

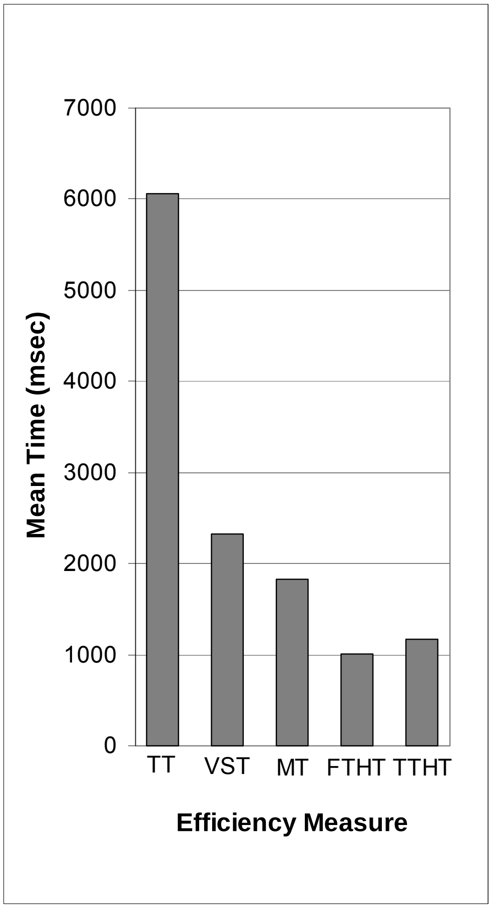


Figure 4.3. Summary of for time-based efficiency measures.

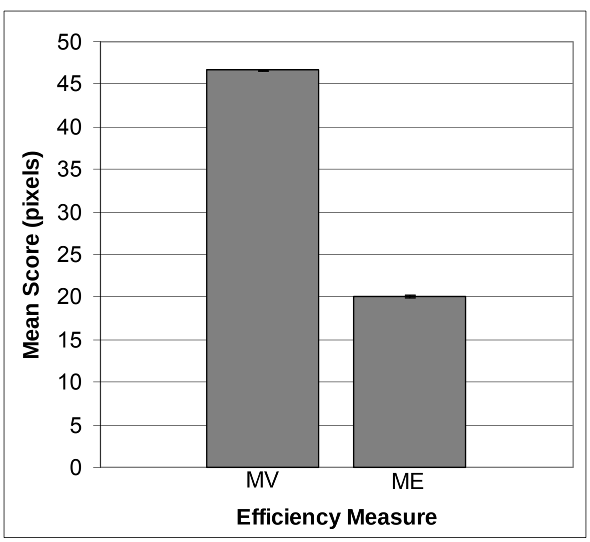


Figure 4.4. Summary of pixel-based efficiency measures.

Analyses: TT, VST, MT, MV, ME

The efficiency measures in the desktop condition, like the handheld, did not meet the assumptions of normality required for linear regression. Specifically, the distributions of the error terms from these dependent variables were not normal and the sample sizes were not large enough to claim the central limit theorem had an impact on the distribution. Transformations were applied to these dependent variables to meet the assumptions, and were kept consistent with the transformations applied in the handheld analyses. These transformations are delineated in Table 4.6.

Forward stepwise linear regressions analyzed the contributions of the predictor variables to the overall variance TT, VST, excluding FTHT and TTHT (which used logistic regression). Two linear models were produced for each efficiency measure, one for the clinically acquired visual factors and one for the non-clinically-acquired visual factors. The distributions for FTHT and TTHT did not meet the assumptions for a regression analyses, even with several transformation attempts. Instead, a logistic regression was applied to the highlight measures. A more detailed description of the analyses and outcomes based on the highlight measures follows the description and discussion of the linear regression analyses for the other efficiency measures.

Table 4.6. Description of efficiency measure transformation and an interpretation of the transformed variable.

Efficiency Measure	Transformation	Interpretation
TT	$1/\sqrt{TT}$	The rate of trial completion: as $1/\sqrt{TT}$ increases, the trial was completed at a faster rate.
VST	$1/\sqrt{VST}$	The rate visual search termination: as $1/\sqrt{VST}$ increases, the visual search completed at a faster rate.
MT	lnMT	The amount of time allocated to movement of the card to the drop pile: a higher lnMT indicates movement time to the icon was slower.
MV	$\sqrt{MV}$	The standard deviation in the distances of the path taken from the task axis from the mean on the movement of the card icon to the drop pile; a higher $\sqrt{MV}$ translates to a higher variability and less efficient movement.
ME	$\sqrt{ME}$	A measure of the average distances of the path taken from the task axis from the mean during the movement of the card icon to the drop pile, a higher $\sqrt{ME}$ the larger the distance and the less efficient the movement.

Results: Clinically-acquired Ocular Factors

Summaries of each regression model generated on the efficiency measures are provided in Table 4.7, and include R^2, R^{2-}adjusted, and the Durbin-Watson statistic for each model. With the exception of MT and ME, the emergent models were all good fits to the data and accounted for between 41.5% to 45% of the variability, based on R^2-adjusted. Durbin-Watson statistics were at acceptable levels (not <1 or > 3, and close to 2). Much of the variability in ME remained unaccounted for in the generated model (~90%), and the R^2-adjusted scores for MT were marginal, at best. It is therefore assumed that other factors, not integrated into this model, were largely driving the measurement of ME. Because the results from ME and MV were also inconclusive in the handheld experiment, they are excluded from further analyses and interpretation in the desktop model. MT is still reported, but discretion is used in the interpretation of related outcomes.

Table 4.7. Clinically-acquired model summary for efficiency measures.

Statistic	**1/√TT**	**1/√VST**	**lnMT**	**√MV**	**√ME**
N	1712	1714	1699	1717	1711
R^2	0.459	0.451	0.308	0.419	0.326
R^2-adjusted	0.452	0.442	0.300	0.415	0.106
Durbin Watson	1.117	1.85	1.386	1.917	2.01

Based on the models generated, predictive equations were generated which can be used to determine the quantitative impact imposed by the factors on efficiency.

1/√TT = .0132 -.001 *SS* -.0005 *ISp* + .00000980 *Trial #* -.0001 *Age* + .0002 *SF12MCS* + .0000483 *SF12PCS* + .0003 *Dexterity* - .0083 *NVA* + .0006 *AMD*Sz*

1/√VST = -.000387 + .00129 *Sz* - .00171 *SS* - .000740 *ISp* - .00176 *AF* + .0000125*Trial #* + .000256 *SF12MCS* + .0000828 *SF12PCS* + .000760 *Dexterity* - .00947 *NVA* + .000510 *AMD*Sz* + .000555 *AMD*AF*

lnMT = 4.87 + .0539 *Sz* + .1062 *Column* - .0008 *Trial #* + .0294 *Age* - .0093 *SF12MCS* + .063 *Dexterity* + .9865 *NVA* - .1668 *AMD* - .064 *AMD*AF*

While the predictive models generated are of practical utility in estimating the actual efficiency measures in certain conditions, it is important to consider the predictive factors excluded from each model, in addition to the standardized Beta values. Table 4.8 provides an overview of the factors included in the model for each measure, B, S.E. of B and B-std. The standardized coefficient is useful in the practical interpretation of the models. B-std values afford the quantitative comparison assessment of the sizable impact each predictor has on the outcome measure, relative to each other (Field 2000).

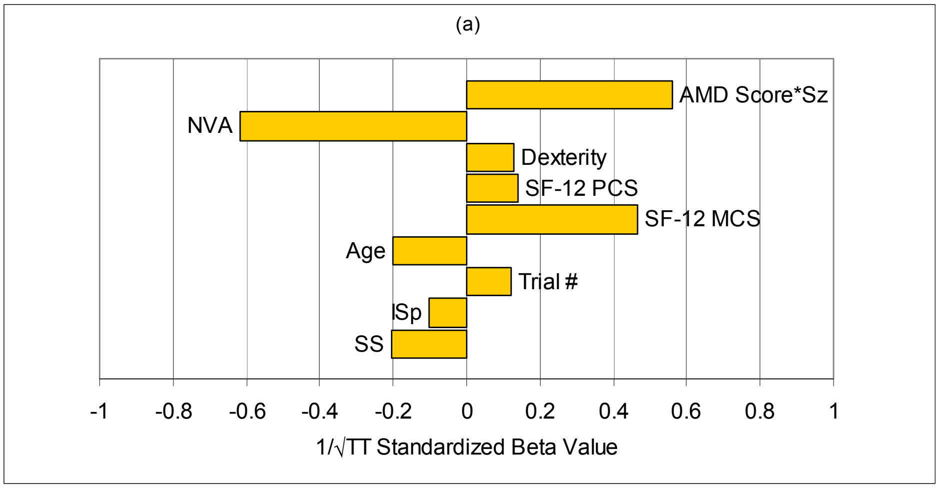
Figures 4.5a-4.5c disclose the relative impact of each predictor variable on the respective models. Variables with bars extending to the left of the reference line at zero imposed a decrease on the model of the outcome variable and bars extending to the right imposed an increase. Because this graph plots the standardized coefficient (B-std), relative comparisons can be made in terms of 'how much more' a predicator influences each model. When interpreting the direction of impact, an increase in $1/\sqrt{TT}$ and $1/\sqrt{VST}$ equates to a faster, improved time, while predicted increases in lnMT and $\sqrt{MV}$ represents slower movement times and indicate more variability in the movement of the mouse on the way to the drop (i.e. diminished efficiency). The figures facilitate the inference of the cumulative effect of the predictors on each outcome measure in relation to each other. In addition, the consideration of different scenarios for various participant abilities and task factors can be approximated to determine the optimal design strategy.

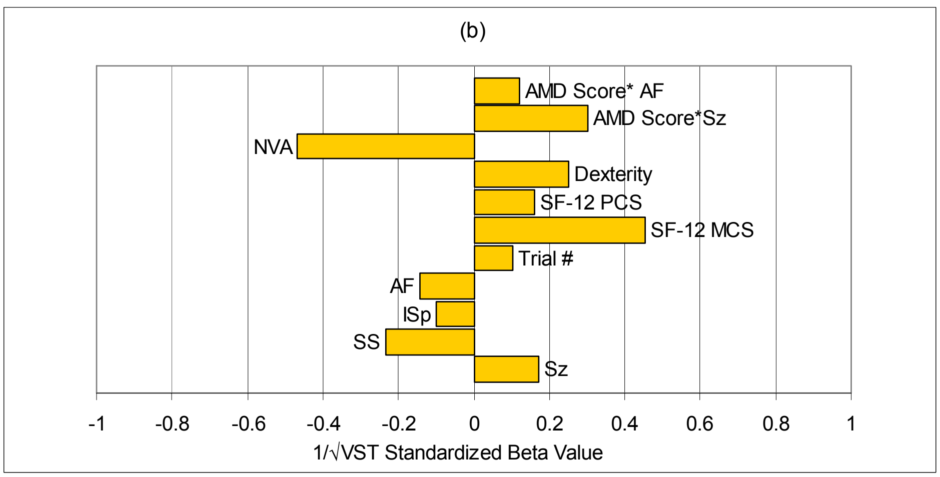
Table 4.8 Summary of predictor variables for efficiency measure regression with clinically-acquired ocular factors (***** indicates that factor was excluded from the model).

Task-related Factors				
Variable		**1/√TT**	**1/√VST**	**lnMT**
Constant	*B*	0.013	-3.88E-04	4.876
	SE	0.003	0.002	0.364
	p	<.001	0.872	<.001
Sz	*B*	*****	0.001	0.054
	SE		<.001	0.020
	B-std		0.170	0.087
SS	*B*	-0.001	-0.002	*****
	SE	<.001	<.001	
	B-std	-0.204	-0.232	
ISp	*B*	-0.001	-0.001	*****
	SE	<.001	<.001	
	B-std	-0.104	-0.100	
AF	*B*	*****	-0.002	*****
	SE		<.001	
	B-std		-0.144	
Column	*B*	*****	*****	0.106
	SE			0.023
	B-std			0.146
Row	*B*	*****	*****	*****
	SE			
	B-std			
Drop Location	*B*	*****	*****	*****
	SE			
	B-std			
Trial #	*B*	9.801E-06	1.252E-05	-0.001
	SE	<.001	<.001	<.001
	B-std	0.120	0.102	-0.080

Table 4.8. continued.

Participant-related Factors				
Variable		**1/√TT**	**1/√VST**	**lnMT**
Age	*B*	-1.00E-04	*****	0.0294
	SE	<.001		0.003
	B-std	-0.202		0.444
SF-12 MCS	*B*	2.00E-04	2.57E-04	-0.009
	SE	<.001	<.001	0.002
	B-std	0.466	0.454	-0.2
SF-12 PCS	*B*	4.84E-05	8.28E-05	*****
	SE	<.001	<.001	
	B-std	0.142	0.161	
Dexterity	*B*	3.00E-04	7.60E-04	0.063
	SE	<.001	<.001	0.014
	B-std	0.129	0.253	0.25
Clinically-acquired Ocular Factors				
Variable		**1/√TT**	**1/√VST**	**lnMT**
NVA	*B*	-8.30E-03	-9.47E-03	0.9865
	SE	0.001	0.001	0.084
	B-std	-0.616	-0.469	0.587
CS	*B*	*****	*****	*****
	SE			
	B-std			
AMD Score	*B*	*****	*****	-0.167
	SE			0.02
	B-std			-0.481
AMD * Sz	*B*	6.00E-04	5.10E-04	*****
	SE	<.001	<.001	
	B-std	0.559	0.303	
AMD * SS	*B*	*****	*****	*****
	SE			
	B-std			
AMD * ISp	*B*	*****	*****	*****
	SE			
	B-std			
AMD * AF	*B*	*****	5.55E-04	-0.064
	SE		<.001	0.016
	B-std		0.121	-0.168





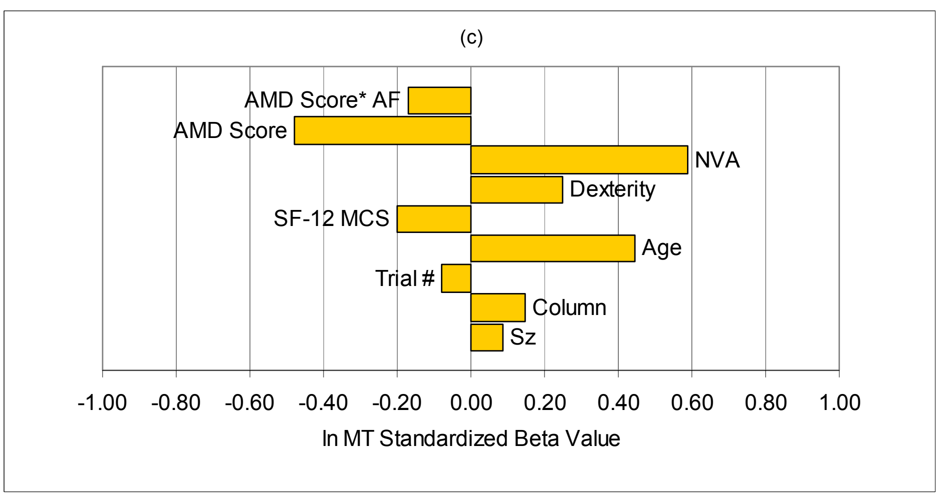


Figure 4.5 a-c. The relative impact of clinically-acquired predictor variables, illustrated via B-std for the accuracy measures.

Results: Non-Clinically-acquired Ocular Measures

The analyses of the efficiency measures were replicated using non-clinically-acquired ocular factors (i.e. VSAT percentile and overall VFQ score) in lieu of the clinically-acquired (CS, NVA, and AMD Score), maintaining the inclusion of the personal and task-related factors. Dependent measures were transformed for consistency with the previous clinical models to meet the assumptions required for linear regression. Linear regression models were generated for TT, VST, and MT. while logistic regression was applied to the target highlight time metrics (TTHT and FTHT).

Table 4.9 presents the non-clinically-acquired ocular factor model summary for all efficiency measures, reporting R^2, R^{2-}adjusted and Durbin Watson statistic for each model. The models for TT and VST presented somewhat good fits of the data, and accounted for 28.1% and 34.6% of the variability in the dataset, respectively. The model for MT does not represent a ‘good fit’ and accounts for less than 14% of the variability in the movement. The results of MT for the non-clinical model while they are provided in this chapter are not regarded as highly informative. This poor fit to the model was also observed MT in the clinically-acquired model.

Table 4.9. Non-clinically-acquired model summary for efficiency measures.

Statistic	$1/\sqrt{TT}$	$1/\sqrt{VST}$	lnMT
N	1715	1715	1701
R^2	0.289	0.353	0.146
R^2-adjusted	0.281	0.346	0.136
Durbin Watson	0.971	1.233	1.138

Based on the models, a linear equation was generated for each efficiency measure, the components summarized in Table 4.10. The predictive equations for modeling $1/\sqrt{TT}$, $1/\sqrt{VST}$, and lnMT, were as follows:

$$1/\sqrt{TT} = .0032 - .001\ SS - .0005\ ISp + .0000102\ Trial\ \# + .0002\ SF12MCS + .0000596\ SF12PCS + .0002\ Dexterity + .0000147\ VFQ*Sz - .0000102\ VSAT*Sz$$

$$1/\sqrt{VST} = -.0017 + .003\ Sz - .0017\ \mathrm{SS} - .0007\ ISp + .0000114\ Trial\ \# + .0002\ SF12MCS + .0005\ Dexterity + .0001\ VFQ - .0000128\ VFQ*Sz$$

$$lnMT = 6.8349 + .0443\ ISp + .1057\ Column - .0015\ Trial\ \# + .0127\ Age - .0138\ SF12MCS - .0063\ SF12PCS + .0422\ Dexterity + .0019\ VSAT*Sz$$

Table 4.10. Summary of predictor variables for efficiency measure regression with non-clinically-acquired ocular factors (***** indicates that the exclusion of that predictor for that model).

Task-related Factors				
Variable		**1/√TT**	**1/√VST**	**lnMT**
Constant	*B*	0.0032	-0.0017	6.8349
	SE	1.00E-03	3.00E-03	0.365
	p	0.014	0.53	<.001
Sz	*B*	*****	0.003	*****
	SE		0.001	
	B-std		0.393	
SS	*B*	-0.001	-0.0017	*****
	SE	<.001	<.001	
	B-std	-0.211	-0.236	
ISp	*B*	-0.0005	-0.0007	0.0443
	SE	<.001	<.001	0.022
	B-std	-0.105	-0.101	0.072
AF	*B*	*****	*****	*****
	SE			
	B-std			
Column	*B*	*****	*****	0.1057
	SE			0.026
	B-std			0.145
Row	*B*	*****	*****	*****
	SE			
	B-std			
Drop Location	*B*	*****	*****	*****
	SE			
	B-std			
Trial #	*B*	1.022E-05	1.137E-05	-0.002
	SE	<.001	<.001	<.001
	B-std	0.125	0.092	-0.144

Table 4.10. continued.

Clinically-acquired Ocular Factors				
Variable		**1/√TT**	**1/√VST**	**lnMT**
VSAT	*B*	*****	*****	*****
	SE			
	B-std			
VFQ	*B*	*****	1.00E-04	*****
	SE		<.001	
	B-std		0.405	
VFQ*Sz	*B*	1.47E-05	-1.28E-05	*****
	SE	<.001	<.001	
	B-std	0.276	-0.16	
VFQ*SS	*B*	*****	*****	*****
	SE			
	B-std			
VFQ*ISp	*B*	*****	*****	*****
	SE			
	B-std			
VFQ*AF	*B*	*****	*****	*****
	SE			
	B-std			
VSAT*Sz	*B*	-1.017E-05	*****	0.0019
	SE	<.001		<.001
	B-std	-0.101		0.15
VSAT*SS	*B*	*****	*****	*****
	SE			
	B-std			
VSAT*ISp	*B*	*****	*****	*****
	SE			
	B-std			
VSAT	*B*	*****	*****	*****
	SE			
	B-std			

Table 4.10. continued.

Participant-related Factors				
Variable		**1/√TT**	**1/√VST**	**lnMT**
Age	*B*	*****	*****	0.0127
	SE			0.003
	B-std			0.19
SF-12 MCS	*B*	2.00E-04	2.00E-04	-0.014
	SE	<.001	<.001	0.002
	B-std	0.426	0.307	-0.293
SF-12 PCS	*B*	5.961E-05	*****	-0.006
	SE	<.001		0.002
	B-std	0.173		-0.146
Dexterity	*B*	2.00E-04	5.00E-04	0.0422
	SE	<.001	<.001	0.013
	B-std	0.113	0.151	0.166

Figure 4.6a-c illustrates the relative impact of each predictor variable in these models using the non-clinically-acquired metrics. Variables with bars extending to the left of the 0 line imposed a decrease on the model of that measure, and to the right imposed an increase. Again, when interpreting the direction of impact, recall that an increase in 1/√TT and 1/√VST equates to a faster, improved time, while predicted increases in lnMT and slower rates of icon movement.

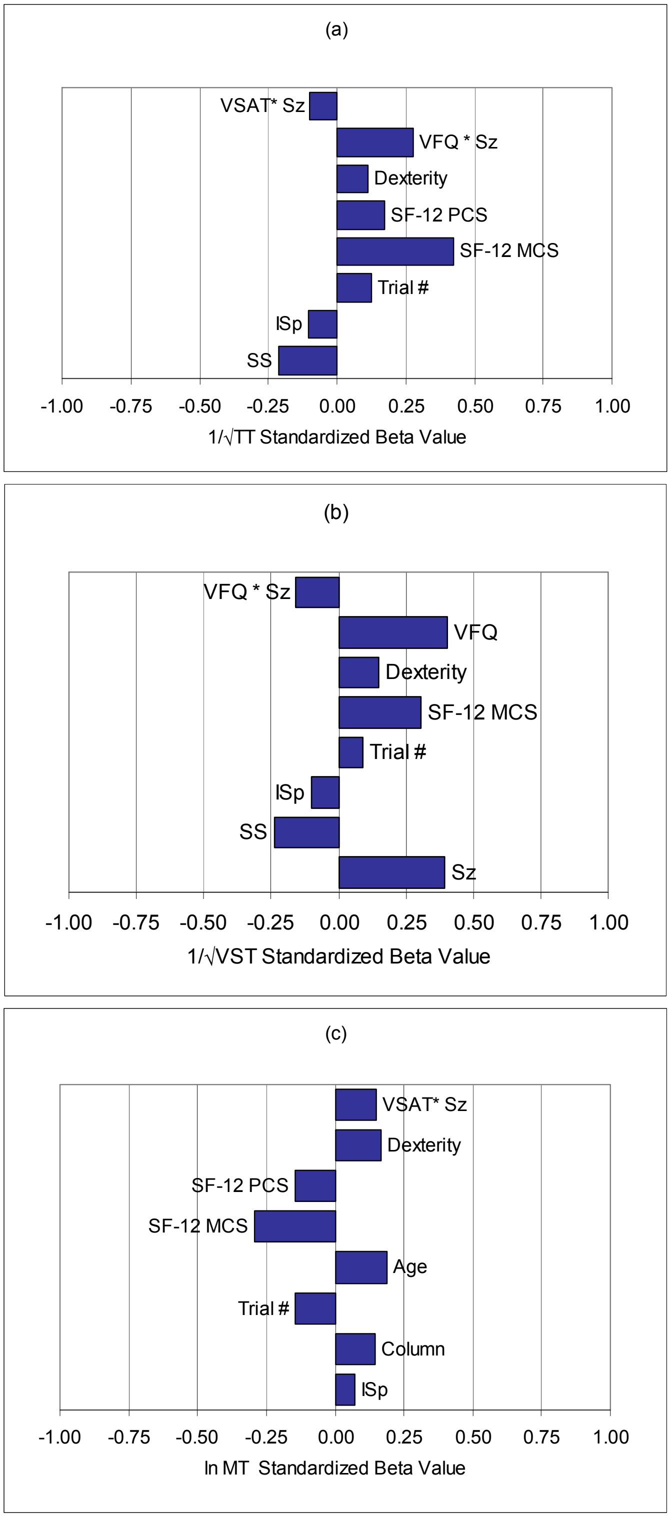


Figure 4.6 a-c. The relative impact of non-clinically-acquired predictor variables, illustrated by B-std.

TT, VST, DD, MT Outcome Summary

Clinically-acquired Models

The outcomes of the regressions on TT, VST, DD and MT, revealed many interesting trends in the participants' interactions with the desktop. This section details these outcomes which will be addressed again at the conclusion of the chapter in terms the hypotheses introduced in Chapters 1 and 2.

Outcome #1: Independent Controlled Interface Variables

Set size and Icon Spacing. Increases in the SS and ISp were prone to trigger slower rates visual search termination (e.g., VST) and longer TT. These effects were small compared to the influence attributed to visual and personal factors (on the order of ½ the impact or more). However, it is interesting to consider that these trends were observed to be main effects across all participants. This shows that there is an eventual point of diminishing returns for these interface variables, which can impact overall task efficiency (TT). Increases in SS and ISP generated slower TT and VST for all the participants, with all other factors being constant.

Icon size. Sz contributed to some interesting effects on the different interaction phases. While the effect of Icon size was not influential on TT across all participants, it demonstrated a minimal degree of influence on both VST and MT. In the model, as Sz increased, VST decreased. However in an opposing influence, increases in the size of the icons posed a tendency to increase MT, suggesting less efficient behavior in the manipulation of the larger icons across the display.

Auditory Feedback. The inclusion of supplemental non-visual cues in the task did not have widespread main effects across all participants; however, there was a negative impact of AF in the model of VST. In this model VST slowed down in the presence of auditory feedback across all participants. There was an interaction nonetheless between AMD Score and AF, which essentially canceled this negative influence of AF presence.

The negative influence of AF on VST therefore only applies to individuals who fall within the control group. This result is particularly compelling when considering that AF does not apply to the visual search portion of the experimental task. This indicates that the presence of AF in the interaction may impose negative carry over effects into other phases of the task, which in turn can distract the users from effectively meeting their goals.

Outcome #2: Clinically-acquired Visual Factors

Near Visual Acuity. NVA was the most influential factor in all three clinical models of efficiency. As NVA worsened (from .1 to 1), a monotonic trend is observed in the model in terms of diminished performance levels. TT, VST and MT are observed to increase – or slow down in the presence of diminished NVA. Previous studies have shown significant influence of NVA and AMD on task interaction in a simple drag and drop task. The emphasis of NVA in the models implicates this aspect of visual function as critical to the efficient completion of the task.

Severity of AMD and Contrast Sensitivity. The impact of AMDScore and CS on the efficiency models was not as pronounced as is was for NVA. CS was not included as a predictor in any of the three models. For AMD score, only one main effect was observed. Increased severity levels influenced decreased MT. As AMD Score increased in severity, the icons were moved to the drop piles at a faster rate. This result is difficult to account for, and furthermore when the interactions of AMD and interface factors are considered. It appears that the interaction AMD * AF also influenced MT. Increases in AMD*AF generated decreased MT as well, which implies the individuals without AMD did not experience the same performance gains attributable to the inclusion of AF. Additional clarification of the movement interaction is needed at a more granular level to more completely appreciate the relationship between MT, ocular disease severity and AF. The analyses of the accuracy measures will provide further insight on this topic. The results for MT are judiciously considered in light of the relatively low R^2-Adjusted level calculated for the model.

Outcome #3: Interactions between Visual and Interface Factors

*AMD Score * Sz*. The relationship between AMD*Sz was included as predictor in the model of VST and TT. As the interaction term increased, the rate of VST and TT were faster in the model. This implies that the benefits of increased Sz were especially pronounced in the population with AMD, and even more so at more pronounced levels of the disease. Furthermore, in terms of TT, Sz did not produce measured benefits across the entire set of participants, only for those with AMD present. This suggests the impact of SS on the overall task is positive. The positive interaction of Sz with AMD also counteracts the decreases in TT and VST observed relative to decrements in NVA.

*AMD Score * AF.* As discussed, AF had a negative influence on the model of VST, responsible for slower search times. However, the influence of AF is negated in the model for those individuals with any level of AMD severity. Furthermore, there is potential in the model for improved VST for those individuals with more severe stages of AMD. Again, this suggests that the AF could have been a potential a source of diversion in the context of task performance for those who could more clearly acquire visual information in the interface. Also, because those trials including AF were grouped together sequentially, this result suggests the possibility of carry over effects of the AF on components of the task other than the ‘drop’ of the card into the pile.

Outcome # 4: Participant-related Factors

Age. The influence of age on the interaction efficiency was present, but weighed most heavily on the movement time. Increases in age were observed, in the model, to influence increased overall time, for slower rates of task completion, and also increased MT for slower relocation of the card icon to the drop pile. While neither result is surprising, the results demonstrate the important consideration of age-related interaction differences, even within a population limited in their age range, such as the target population with AMD. Differences in performance, abilities, and coping skills are highly correlated with age, and differences are observed at the different age groups within the older adult population. In other words, while the investigation considered only individuals over 55, the segment of the general population for which AMD is prevalent, there are measurable differences between participants behaviors at the “young old” (<65), older adults (65-75) and the “old-old” population segments (>75) (Rogers 1997; Smith, Sharit and Czaja 1999; American Foundation for the Blind 2004). The influence of age on the desktop interaction was not fully captured by the constructs of ocular health, SF-12, or the test of manual dexterity. Underlying nuances of age that affect the interaction, unaccounted for in the constructs of measured by the visual factors, SF-12, and manual dexterity.

Manual Dexterity. Improvements to manual dexterity, attributed to increasing average scores to the Purdue Pegboard test, were shown in the model to trigger faster TT, VST, but slower MT. This suggests that the impact of dexterity deviates between the different phases of task performance. In terms of global task performance, the overall effect of improved levels of manual dexterity is logical, and its influence over other factors is not inflated. While the effect on MT is not entirely logical, the investigation of task accuracy has potential to reveal the reasoning that guides this outcome. Also, the low R^2-squared value for MT model should be remembered in weighing the levity of this outcome.

Mental & Physical Health. The SF-12 subscales, PCS and MCS, were both included predictors for the clinical models of task efficiency. The effect of both was not surprising. PCS and MCS were shown to influence faster TT, VST, while only MCS influenced faster MT. Perhaps more interesting is that the influence of MCS was on the same order of magnitude as the ocular factors in TT and VST. The effects imposed were second only to NVA in the VST model and closely followed the amount of impact both NVA and AMD*Sz had in the model of TT.

Outcome #5 Other Interface Features

Learning. Cases with a Trial# positioned later in the context of the 162 trials had a small influence on faster TT, VST and MT. The influence of learning in this task, however, contributed roughly one-third of the influence that the visual factors had over VST and TT. This demonstrates that while individuals can make small performance gains with repetition in interface use, learning cannot completely compensate for the limited bandwidth of the visual sensory channel in a dynamic visual interface.

Card Icon & Drop Pile Locations. Of the predictors indicative of card icon and drop pile location (row, column, and drop location) the column of the icon was included as a predictor for efficiency. Columns further to the right of the display, logically, impose increases in the movement time. The location of targets on the display did not influence the visual search time, or the overall trial time.

Non-Clinically-acquired Models

Outcome#1: Differences between Clinical and Non-Clinical Models

Table 4.11 summarizes the consistencies and inconsistencies between the models generated for the efficiency measures using the clinical versus the non-clinical factors. While the complete set of factors included in each class of models differed, there were no cases in which the effect of a predictor variable was observed to behave in a conflicting way between the clinical and non-clinical models. The influence of the interface variables was consistent between the models of TT, but deviated slightly in the model of VST (AF had a measured influence in the clinical model), and MT.

Table 4.11. Deviation of predictor variables between clinical and non-clinical models.

Model	Consistent	Unique to Clinical Models	Unique to Non-Clinical Models
1/√TT	Dexterity SF-12 PCS SF-12 MCS Trial # ISp SS	AMD*Sz NVA Age	VSAT*Sz VFQ*Sz
1/√VST	Dexterity SF-12 MCS ISp SS Sz	AMD*AF AMD*Sz NVA SF-12 PCS AF	VFQ VFQ*Sz Age
LnMT	Dexterity SF-12 MCS Age	AMD Score AMD*AF NVA Drop Location Sz	VSAT*Sz SF-12 PCS Trial # Column ISp

Outcome #2: Independent Controlled Interface Variables

Set Size and Icon Spacing. The influence of SS and ISp was consistent with the observed effects in the clinical models of efficiency, with an additional influence of ISp on MT. Increases to both ISp and SS in the models led to slower TT and VST. Increases in ISp were also observed in the non-clinical models to generate a decrease in MT. Overall the extent of the influence of ISp on the models was small when compared to the B-std of the other factors, particularly the visual factors; and there were no interactions between the visual factors and ISp. The changes in SS, while they had a greater level of influence on TT and VST, were not shown to interact sufficiently with the visual factors for inclusion in the models.

Icon Size. A main effect of Icon size was observed in the model for VST. Visual search was likely to be terminated more quickly for the larger Icon sizes.

Outcome #3 Non-Clinically-acquired Visual Factors

VFQ. There was little observed influence in terms of main effects of non-clinical visual factors. The only factor to be included in a model was the VFQ overall average in the prediction of VST. As the VFQ increased, the model reflected faster VST. This was the most significant factor influencing VST, followed closely by Sz and SF-12 MCS.

Outcome #4: Interactions between VSAT, VFQ and Interface Variables

Interactions with Icon Size. The only interactions observed in the non-clinical models were attributed to the relationship between VFQ and VSAT with Sz. The VSAT*Sz interaction demonstrated that increases in this term led to slower TT and slower MT. While the magnitude of this influence was not great, it suggests an important trend underlying the efficiency models. Increased icon size may not as be helpful in the improvement of overall efficiency (TT) for individuals with high VSAT score. The results of TT maintain that everyone benefits at least a little from the increased icon size, but those with lower levels of Visual attention will benefit most.

VFQ and Sz, however, demonstrated interaction effects opposite in direction to that observed for the VSAT*Sz relationship for TT. An included predictor for the model of TT and VST, the results were not consistent. An increase in the value of the VFQ*Sz term generated faster TT. This can be interpreted to suggest that improvements to a user's perceptions of visual function and daily activity were indicative of better TT in the presence of the larger icons.

Outcome #5: Participant-related Factors

Age. Age was observed to influence slower MT. Increases in age were modeled to induce slower MT, consistent with the clinical model of the desktop, and not unexpected.

Manual Dexterity. The impact of manual dexterity in the three efficiency models was consistent with the results of the clinical models. Better dexterity influenced faster TT and VST, but slower MT. Rationalization for the negative bearing of better manual dexterity on MT will be considered in the analyses of the accuracy measures, in consideration of a speed accuracy trade-off. However, the relatively low R^2-adjusted value calculated for the MT model cannot be ignored in this interpretation.

Mental & Physical Heath. The SF-12 PCS and MCS components, as ratings increased, revealed TT, MT, and VST (MCS only). The degree of influence had these factors over the models was substantial. While interface interventions were not considered in terms of overall mental and physical health, their presence in the model contributes to a more realistic representation of the interaction and more highly generalizable conclusions.

<u>Outcome #6 Other interface features</u>

Learning. A small learning effect was observed on all three efficiency measures. At later trials (larger Trial#) TT, VST, and MT were modeled to be faster than during earlier trials. However, the relative impact of learning in this task was small. The impact of visual dysfunction and personal factors were more dominant in their influence over task efficiency than familiarity with the task and interface through repetition.

Icon Location. The only result to emerge in terms of location of the icons or drop piles was the effect of the Column of the target icon. This result was not surprising, but the inclusion of the term served to strengthen the model. The further the column from the drop piles, the larger the movement time, as reflected in the linear regression model. The impact of column replicated what was observed in the clinical model.

Analyses: TTHT & FTHT

As stated, the distribution of the participant performance on both FTHT and TTHT were not suited for regression analyses, despite several transformation attempts. Instead, logistic regression was used to identify which factors most impacted the probability of each highlight measure exceeding a predetermined threshold value, as in Chapter 3, for the handheld analyses. These threshold values were derived by examining the distribution of the participant highlight scores, and designating a cut point at the 85^{th} percentile. This cut point was chosen for consistency with the handheld analyses. The outcome of the logistic regression can designate how the predictors influence the probability that a target highlight time be classified above the 85^{th} percentile. The predictor variables for these logistic regressions were consistent with those used the regression models generated for the other efficiency measures. The 85^{th} percentile for TTHT was calculated to be 1713 msec, based on average participant performance, this constituted .26% of a typical trial time. The 85^{th} percentile for FTHT was identified as 1572 msec, or .24% of the mean trial time.

Results: Clinically-Acquired Visual Factors

Forward-stepwise logistic regressions, based on the likelihood ratio statistic, were applied to both FTHT and TTHT for the models employing the clinically-acquired measures, consistent with the analyses used in the handheld chapter. The logistic regression produced valid models for predicting the likelihood that either FTHT or TTHT would be in excess of the 85^{th} percentile for this population sample (e.g., the longest highlight times).

The Hosmer and Lemeshow Goodness of Fit assessment (HL test) was used to assess the null hypotheses that there was no significant difference between the observed and predicted values for the dependent variable for each model. The HL was test not significant for the model of TTHT ($p = .271$), but was significant for the model of FTHT ($p =.033$). This indicates the model generated for TTHT was a good fit, but that the model generated for FTHT was not. Furthermore, for this sample population the TTHT model was correct in its categorization of 86.5% of the cases and the FTHT model was accurate in 85.2% of its classifications. For FTHT, the discrepancy between the HL test and the accurate classification may be a result of sample size.

The coefficients, test statistics and significance levels are described by Table 4.12. Figure 4.7 illustrates the magnitude of the impact of each predictor in each model via Exp (B) or ‘change in odds’ for the outcome measure. Exp (B) values of less than one influence a decrease in the probability of the dependent variable and Exp (B) values that are greater or equal to one increase the likelihood of the outcome. As the bar extends further from 1, in either direction, the impact on the change in odds is greater.

Table 4.12. Significant variables associated with logistic regressions on FTHT and TTHT using clinically-acquired measures (***** indicates exclusion of that predictor from the model).

TTHT				
Variables	**Coefficients (B)**	**Wald Statistic**	**p**	**Exp (B)**
Constant	-2.670	9.199	0.002	0.069
Sz	*****			
SS	*****			
ISp	*****			
AF	*****			
Column	*****			
Row	*****			
Drop Location	0.268	6.239	0.012	1.307
Trial #	*****			
Age	*****			
SF-12 MCS	-0.059	22.309	0.000	0.943
SF-12 PCS	*****			
Dexterity	0.257	14.352	0.000	1.293
NVA	1.661	10.843	0.001	5.265
CS	*****			
AMD Score	*****			
AMD Score * Sz	*****			
AMD Score*SS	*****			
AMD Score* ISp	*****			
AMD Score* AF	-0.309	6.937	0.008	0.734

Table 4.12. continued.

FTHT				
Variables	Coefficients (B)	Wald Statistic	p	Exp (B)
Constant	-6.913	23.054	<.001	0.001
Sz	0.356	6.325	0.012	1.427
SS	*****			
ISp	*****			
AF	*****			
Column	*****			
Row	*****			
Drop Location	*****			
Trial #	*****			
Age	*****			
SF-12 MCS	-0.064	25.685	<.001	0.938
SF-12 PCS	*****			
Dexterity	0.164	4.580	0.032	1.178
NVA	4.209	17.711	<.001	67.283
CS	0.141	10.778	0.001	1.151
AMD Score	*****			
AMD Score * Sz	*****			
AMD Score*SS	*****			
AMD Score* ISp	*****			
AMD Score* AF	*****			

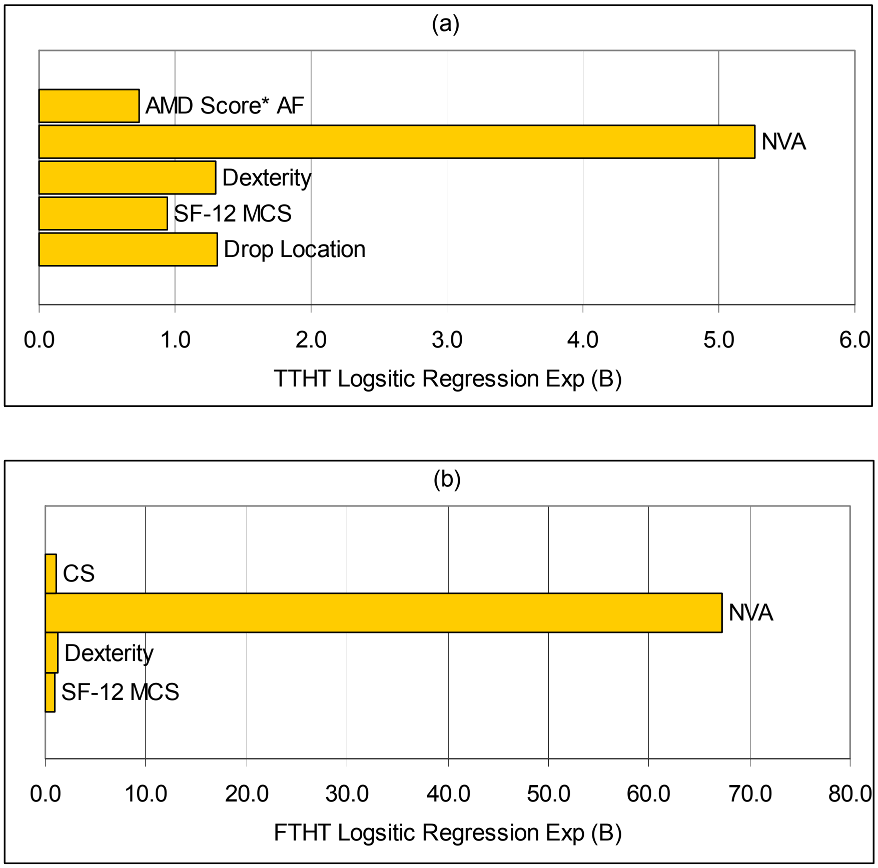


Figure 4.7. Illustration of the relative impact of TTHT and FTHT on the probability of highlight times in excess of the 85th percentile with clinically-acquired predictors.

Results: Non-Clinically-Acquired Visual Factors

Forward-stepwise logistic regressions, based on the likelihood ratio statistic, were applied to both FTHT and TTHT using the non-clinically-acquired factors, personal and task-related factors for consistency with the other models. The HL for TTHT logistic regression was not significant (p= .484) and correctly predicted the outcome in the sample population 86.5% of the time indicating a good fit and model. For FTHT, HL was shown not to be significant (p = .455), and 85.2% of the cases in the sample population were correctly classified by the model. Table 4.13 describes the coefficients, test statistics and significance. Additionally, Figure 4.8 illustrates the magnitude of the impact of each predictor for each model with Exp($\tilde{B}$)

Table 4.13. Significant variables associated with logistic regressions on TTHT and FTHT using non-clinically-acquired measures (***** indicates models predictors exclusion from a model).

TTHT				
Variables	**Coefficients (B)**	**Wald Statistic**	**p**	**Exp (B)**
Constant	-1.260	2.478	0.115	0.284
Sz	*****			
SS	*****			
ISp	*****			
AF	-2.908	5.995	0.014	0.055
Column	*****			
Row	*****			
Drop Location	0.289	7.349	0.007	1.335
Trial #	*****			
Age	*****			
SF-12 MCS	-0.053	21.107	<0.001	0.949
SF-12 PCS	*****			
Dexterity	0.146	5.280	0.022	1.157
NVA	*****			
CS	*****			
VSAT	*****			
VFQ	*****			
VSAT*Sz	*****			
VSAT*SS	*****			
VSAT*ISp	*****			
VSAT*AF	0.029	4.433	0.035	1.030
VFQ*Sz	*****			
VFQ*SS	*****			
VFQ*ISp	*****			
VFQ*AF	*****			

Table 4.13. continued.

FTHT				
Variables	**Coefficients (B)**	**Wald Statistic**	**p**	**Exp (B)**
Constant	-2.513	9.372	0.002	0.081
Sz	0.321	5.362	0.021	1.379
SS	*****			
ISp	*****			
AF	*****			
Column	*****			
Row	*****			
Drop Location	0.238	5.411	0.020	1.269
Trial #	*****			
Age	*****			
SF-12 MCS	-0.055	23.674	<.001	0.947
SF-12 PCS	*****			
Dexterity	0.206	12.638	<.001	1.228
NVA	*****			
CS	*****			
VSAT	*****			
VFQ	*****			
VSAT*Sz	*****			
VSAT*SS	*****			
VSAT*ISp	*****			
VSAT*AF	*****			
VFQ*Sz	*****			
VFQ*SS	*****			
VFQ*ISp	*****			
VFQ*AF	*****			

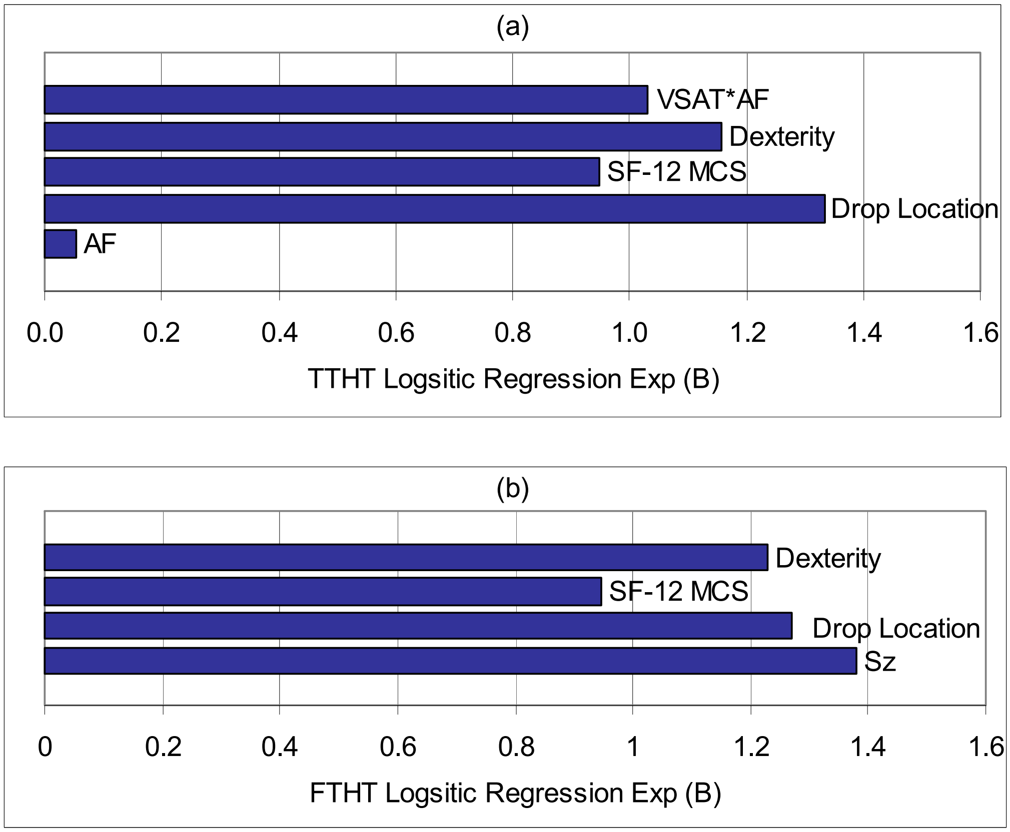


Figure 4.8. Illustration of the relative impact of TTHT and FTHT on the probability of highlight times in excess of the 85th percentile with non-clinically-acquired predictors.

TTHT FTHT Outcome Summary

Clinically-Acquired Visual Factors

Outcome # 1: Clinically-acquired visual factors

Near Visual Acuity. In the models for TTHT and FTHT, NVA was the most dominant force, influencing substantial increases in the probability for longer times on both highlight time metrics. As near vision degraded (the NVA value approached 1), the probability for TTHT and FTHT to exceed the 85^{th} percentile demonstrated sizable gains. This suggests that participants with worse NVA took longer to ensure their icon was in the correct position for release into the card pile, even on the last approach to the card.

Contrast Sensitivity. Contrast sensitivity was an included predictor in the model of FTHT, but not on TTHT. This implies that during the final approach of the icon to the drop pile, improvements in contrast sensitivity would generate a longer highlight time – that is the participants with better contrast sensitivity were slower in their release the mouse button when the icon was in the correct position for a drop. The scale of the quantity of this influence was substantially small in comparison to the impact of NVA on FTHT. In addition, CS was not included as a predictor in any of the linear regression models for the other efficiency measures.

Outcome #2: Interactions

*AMD*AF.* While none of the controlled interface factors (Sz, ISp, SS, AF) demonstrated main effects on the highlight measures across all participants, the AMD*AF interaction was found to contribute to shorter TTHT. This implies that for individuals with AMD, as the severity level of the disease increased in the presence of AF, they were faster with the drop portion of the task. This implies that their accuracy was affected, as a shorter TTHT means fewer approaches to the card pile, and fewer accidental drops. This does not suggest that AF detracts from the movement/drop component of the task for the visually healthy population, as it did in the model of VST.

Outcome # 3 Personal Factors

Manual Dexterity. Increases in manual dexterity were shown to increase the likelihood for TTHT to be classified above the 85th percentile, and had a similar impact on the FTHT model. Because this result is counterintuitive, the analyses of the accuracy measures will provide the insight necessary to explain this trend. However, the effect on TT, VST and MT was more logical, so this may reflect a potential tradeoff between speed and accuracy for those with higher levels of manual dexterity. This will be considered in the analysis of the accuracy metrics.

Mental Health. Mental health, as measured by the SF-12 MCS, was an included predictor in both FTHT and TTHT models. In these models, the likelihood for the times to exceed the 85th percentile was reduced in the presence of increasing levels of self-rated mental health. While an intuitive result, the size of the effect is quite small in comparison to the ocular factors.

Outcome #4: Other interface features

Drop location was the only supplemental interface characteristic to emerge in the highlight models. Drop location influenced TTHT. The drop piles lower on the display imposed a slight increase on TTHT, but without influence over FTHT for locations lower on the display.

Non-Clinically-acquired Visual Factors

Outcome #1 Comparison with Clinically-acquired Models

Several similarities emerged between the clinical and non-clinical models for highlight times. The relative impact of Dexterity and SF-12 MCS were consistent, as was the effect of Drop location. Drop location was also included, in the non-clinical models as influential on FTHT. Increased FTHT was more apt to exceed the 85th percentile in the release of cards into card piles located lower on the display. Finally, AF and Sz demonstrated influence on the models for TTHT and FTHT, respectively.

Outcome #2 Controlled Interface Factors

Auditory Feedback. In the non-clinical model of TTHT, AF had the largest influence on the highlight time, drastically decreasing the likelihood of the longer times across all participants. That said, the interaction between VSAT*AF slightly increasing TTHT, suggests that the participants with worse visual attention skills benefit more from the supplemental non-visual feedback, during the drop component of the task.

Icon Size. Increases in the size of the icons demonstrated an affect on FTHT. Larger icons were prone to cause longer FTHT, which suggests that there is more difficulty in alignment of larger icons over the pile for correct drops, and the smaller icons are more easily positioned prior to the final release of the mouse button.

Outcome #3 Interface Variables and Visual Factor Interaction

*VSAT*AF.* The logistic regressions founded in the non-clinical visual factors were not as telling about the impact of visual profile on this component of the task. The only visual factor to bear influence on highlight time was the VSAT*AF interaction, which implied a higher likelihood for longer TTHT in the presence of AF, amplified in the presence of high levels of visual attention. The relative impact of this increase in likelihood was slight compared with the other included factors.

4.3 Accuracy

General Summary

The mean scores for each accuracy measure, as summarized in Table 4.14, reinforced the high level of success participants experienced with their interactions on the desktop. The low error rate in this study is not surprising, as other investigations with similar populations (and no time limit on the task) showed similar indications of task efficacy (Jacko, Moloney, Kongnakorn et al. 2005). Figure 4.9. provides additional insight into those accuracy measures summarized in Table 4.14, highlighting the frequency with which each error occurred across participants. The errors principally occurred with a frequency of 1 or not at all. Accordingly, logistic regression was applied to the accuracy measures to determine which predictor variables increased the probability of one or more errors occurring in a single trial (Fahrmeir and Tutz 1994; Neter, Kutner, Nachtsheim and Wasserman 1996).

Table 4.14. Summary of accuracy measures of desktop interaction, (abbreviations for each measure appear in italics).

Accuracy Measure (n=1922)	
Number of missed opportunities: Icon Acquisition *IA*	Mean (Std. Error): .50 (.023) Median: <.001 Minimum: 0 Maximum: 12
Number of missed opportunities: *Over No Drop* *OND*	Mean (Std. Error): .50 (.020) Median: <.001 Minimum: 0 Maximum: 7
Number of accidental drops *AD*	Mean (Std. Error): .17(.013) Median: <.001 Minimum: 0 Maximum: 7
Number of task axis crossings: Icon dragging *TX*	Mean (Std. Error): 1.78 (.042) Median: 1 Minimum: 0 Maximum: 22
Number of movement direction changes: Icon dragging *MDC*	Mean (Std. Error): 3.08 (.086) Median: 3.00 Minimum: 0 Maximum: 114

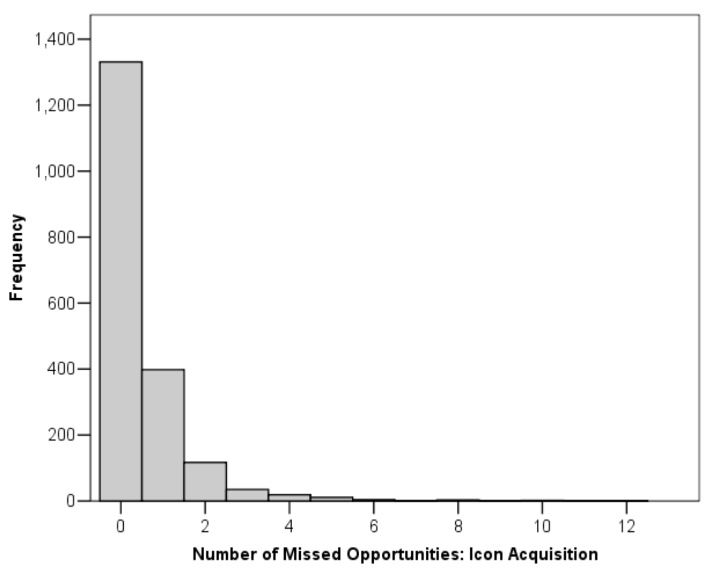


Figure 4.9. Frequency distribution of accuracy measures

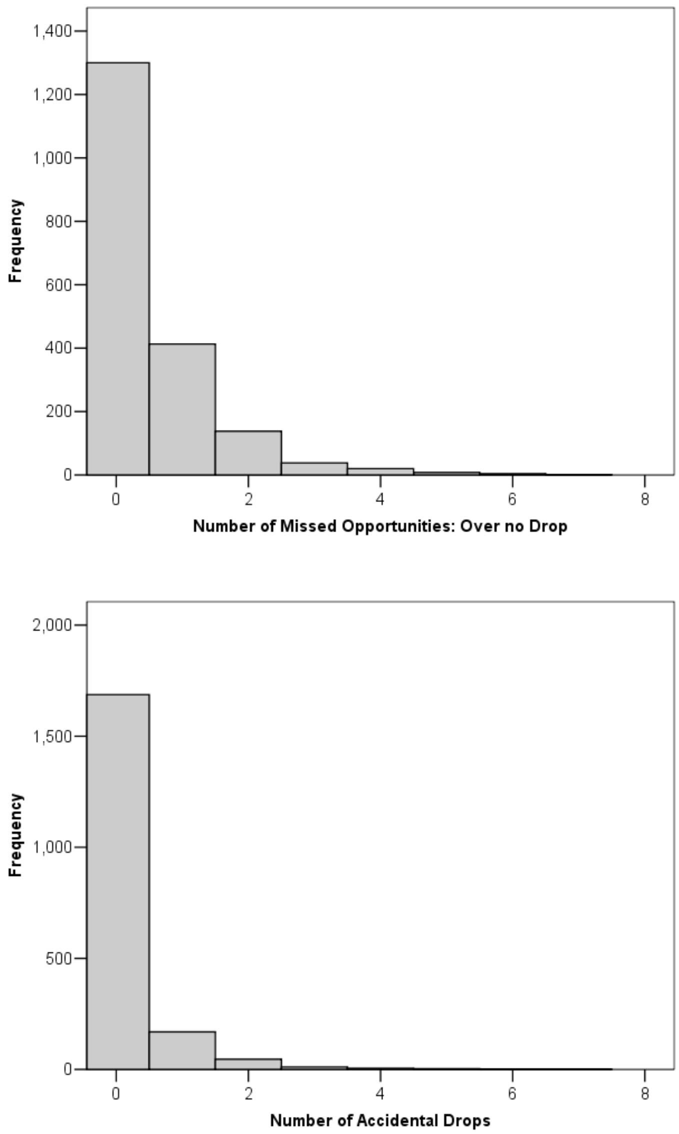
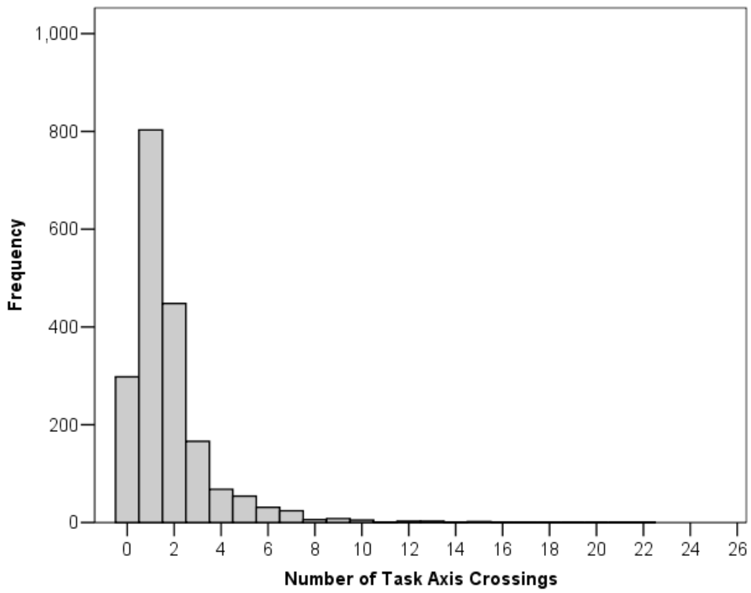


Figure 4.9. continued.



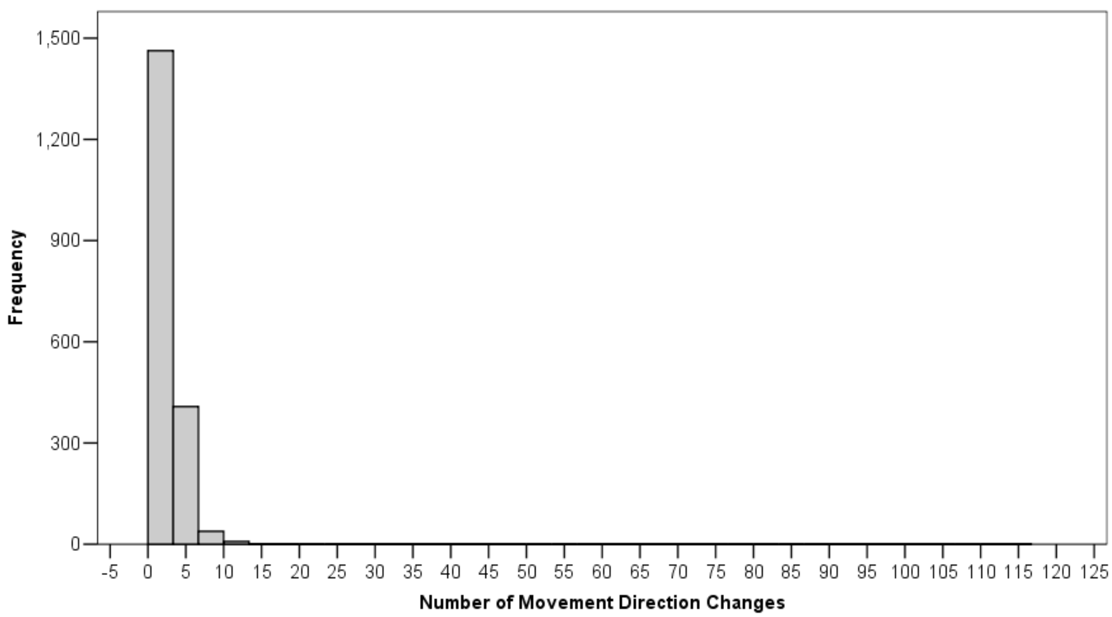


Figure 4.9. continued.

Analyses: Accuracy Measures

As demonstrated by Figure 4.9, the frequency with which accuracy errors occurred during the experimental task was low. Instead of evaluating these accuracy measures as continuous variables using linear regression, the accuracy measures were coded as dichotomous variables (e.g., 0 in those cases where no errors were committed and 1 in cases where 1 or more of that type of error was committed). Using these dichotomous variables, logistic regression models examined the impact of the predictor variables on the likelihood of committing an accuracy error. The predictors considered for inclusion were kept consistent with those utilized in the efficiency analyses. Likewise, stepwise logistic regression was applied to generate the models, using the likelihood ratio method. As with the logistic regressions on the highlight times, the HL test and the percentage of cases correctly classified were used to evaluate the usefulness of the resulting models.

Results: Clinically-acquired Ocular Factors

Table 4.15 summarizes the outcomes of the logistic regression models for IA, OND, AD, TX, and MDC errors based on the clinically-acquired ocular predictors. All models demonstrated a failure to reject the null hypotheses of the HL test ($p \geq 0.8$); the predicted values not significantly different from the observed dependent variables. In addition, percentage of correctly classified cases for each model was at an acceptable level (between 70-97.4%), especially considering the inherent variability of this population. The coefficients (B and exp (B)), test statistics and significance for each efficiency measure are presented in Table 4.16. Figure 4.10

reveals the magnitude of impact each included predictor variable had on the respective models. A graphic is not provided for TX, as only one predictor (AMD*SS) was included in the model.

Table 4.15. Assessment of logistic regression models for accuracy measures.

Variable	HL Goodness of fit test	% Cases Correctly Classified
IA	0.08	71.0%
OND	0.11	70.7%
AD	0.07	86.6%
TX	0.58	83.8%
MDC	0.53	97.4%

Table 4.16. Summary of logistic regression outcomes for accuracy measures using clinically-acquired measures (***** indicates predictors excluded from the models).

IA				
Variables	**Coefficients (B)**	**Wald Statistic**	**p**	**Exp (B)**
Constant	0.45	1.04	0.31	1.57
Sz	-0.45	17.56	<.001	0.64
SS	*****			
ISp	*****			
AF	*****			
Column	*****			
Row	*****			
Drop Location	*****			
Trial #	*****			
Age	*****			
SF-12 MCS	-0.02	6.56	0.01	0.98
SF-12 PCS	*****			
Dexterity	*****			
NVA	*****			
CS	*****			
AMD Score	0.32	31.36	<.001	1.38
AMD Score*Sz	*****			
AMD Score*SS	*****			
AMD Score* ISp	*****			
AMD Score* AF	*****			

Table 4.16. continued.

OND				
Variables	**Coefficients (B)**	**Wald Statistic**	**p**	**Exp (B)**
Constant	3.88	17.67	<.001	48.44
Sz	-0.45	18.81	<.001	0.64
SS	*****			
ISp	-.29	8.12	<.001	0.75
AF	*****			
Column	-0.46			
Row	*****	13.18	<.001	0.63
Drop Location	*****			
Trial #	-3.83E-03			
Age	-0.03	5.31	0.02	1.00
SF-12 MCS	*****	9.62	<.001	0.97
SF-12 PCS	*****			
Dexterity	*****			
NVA	*****			
CS	*****			
AMD Score	*****			
AMD Score*Sz	*****			
AMD Score*SS	*****			
AMD Score* ISp	*****			
AMD Score* AF	*****			

Table 16. continued

AD				
Variables	**Coefficients (B)**	**Wald Statistic**	**p**	**Exp (B)**
Constant	10.33	11.14	<.001	30592.49
Sz	-0.36	4.27	.04	0.70
SS	*****			
ISp	*****			
AF	*****			
Column	*****			
Row	*****			
Drop Location	*****			
Trial #	*****			
Age	-0.07	6.40	0.01	0.93
SF-12 MCS	-0.08	34.38	<.001	0.93
SF-12 PCS	*****			
Dexterity	-0.32	10.47	<.001	0.73
NVA	1.93	8.03	<.001	6.91
CS	*****			
AMD Score	*****			
AMD Score*Sz	-0.20	8.09	<.001	0.82
AMD Score*SS	*****			
AMD Score* ISp	*****			
AMD Score* AF	*****			

Table 4.16. continued.

TX				
Variables	**Coefficients (B)**	**Wald Statistic**	**p**	**Exp (B)**
Constant	1.88	172.07	<.001	6.53
Sz	*****			
SS	*****			
ISp	*****			
AF	*****			
Column	*****			
Row	*****			
Drop Location	*****			
Trial #	*****			
Age	*****			
SF-12 MCS	*****			
SF-12 PCS	*****			
Dexterity	*****			
NVA	*****			
CS	*****			
AMD Score	*****			
AMD Score*Sz	*****			
AMD Score*SS	-0.07	6.19	0.01	0.93
AMD Score* ISp	*****			
AMD Score* AF	*****			

Table 4.16. continued.

MDC				
Variables	**Coefficients (B)**	**Wald Statistic**	**p**	**Exp (B)**
Constant	8.20	39.92	<.001	3.63E+03
Sz	-0.85	5.83	0.02	0.43
SS	-1.40	15.33	0.00	0.25
ISp	*****			
AF	*****			
Column	1.68	8.82	<.001	5.37
Row	*****			
Drop Location	*****			
Trial #	*****			
Age	*****			
SF-12 MCS	*****			
SF-12 PCS	*****			
Dexterity	*****			
NVA	*****			
CS	*****			
AMD Score	*****			
AMD Score*Sz	*****			
AMD Score*SS	*****			
AMD Score* ISp	*****			
AMD Score* AF	*****			

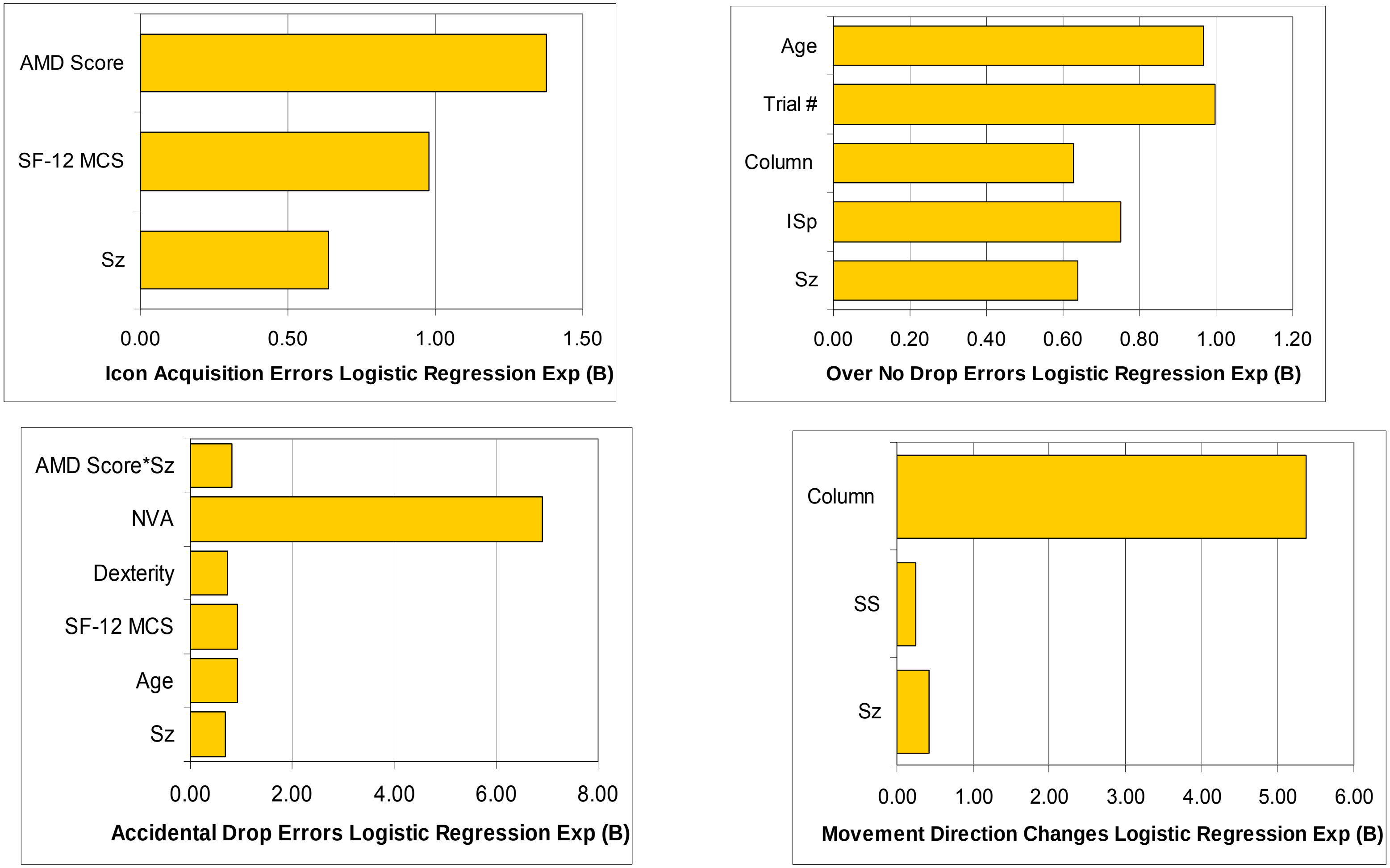


Figure 4.10. Illustration of relative impact of the predictor variables (clinically-acquired) on the likelihood of each error occurring at least once, Exp (B.)

Results: Non-Clinically-Acquired Ocular Factors

Table 4.17 summarizes the assessments of the logistic regression models for IA, OND, AD, TX, and MDC using the non-clinically-acquired ocular factors. All models resulted in a failure to reject the null hypotheses of the HL test. The predicted values therefore were not significantly different from the observed dependent variables. In addition, the percentages of cases correctly classified by each model were at acceptable levels (69-97% accuracy). The coefficients (B and exp (B)), test statistics and significance for each efficiency measure are presented in Table 4.18. Figure 4.11 illustrates the magnitude of impact of each predictor variable for each of the accuracy measures the model via exp (B). The graphic for TX is not provided, as only a single predictor was included in that model (VFQ*Sz).

Table 4.17. Logistic regression model assessment for non-clinical ocular factors and accuracy measures.

Variable	HL Goodness of fit test	% Cases Correctly Classified
IA	0.18	70.4%
OND	0.11	69.3%
AD	0.36	86.6%
TX	0.21	83.8%
MDC	0.53	97.4%

Table 4.18. Summary of logistic regression outcomes for accuracy measures using non-clinically-acquired measures.

IA				
Variables	**Coefficients (B)**	**Wald Statistic**	**p**	**Exp (B)**
Constant	0.57	2.13	0.14	1.78
Sz	*****			
SS	*****			
ISp	*****			
AF	*****			
Column	*****			
Row	*****			
Drop Location	*****			
Trial #	*****			
Age	*****			
SF-12 MCS	*****			
SF-12 PCS	*****			
Dexterity	-0.08	4.10	0.04	0.92
VFQ	*****			
VSAT	*****			
VFQ*Sz	*****			
VFQ*SS	*****			
VFQ*ISP	*****			
VFQ*AF	*****			
VSAT*Sz	-0.01	33.17	<.001	0.99
VSAT*SS	*****			
VSAT*ISP	*****			
VSAT*AF	*****			

Table 4.18. continued.

OND				
Variables	**Coefficients (B)**	**Wald Statistic**	**p**	**Exp (B)**
Constant	3.69	15.92	<.001	40.19
Sz	-0.47	20.35	<.001	0.63
SS	*****			
ISp	*****			
AF	*****			
Column	-0.45	12.26	<.001	0.64
Row	*****			
Drop Location	*****			
Trial #	*****			
Age	-0.04	11.02	<.001	0.96
SF-12 MCS	*****			
SF-12 PCS	*****			
Dexterity	*****			
VFQ	*****			
VSAT	*****			
VFQ*Sz	*****			
VFQ*SS	*****			
VFQ*ISP	-3.02E-03	8.39	<.001	1.00
VFQ*AF	*****			
VSAT*Sz	*****			
VSAT*SS	*****			
VSAT*ISP	*****			
VSAT*AF	*****			

Table 4.18. continued.

AD				
Variables	**Coefficients (B)**	**Wald Statistic**	**p**	**Exp (B)**
Constant	16.53	19.85	<.001	1.51E+07
Sz	-0.75	3.78	0.05	0.47
SS	*****			
ISp	*****			
AF	*****			
Column	*****			
Row	*****			
Drop Location	*****			
Trial #	*****			
Age	-0.13	18.70	<.001	0.88
SF-12 MCS	-0.08	33.79	<.001	0.92
SF-12 PCS	*****			
Dexterity	-0.44	11.43	<.001	0.64
VFQ	*****			
VSAT	*****			
VFQ*Sz	1.36E-03	0.09	0.77	1.00
VFQ*SS	*****			
VFQ*ISP	0.01	5.48	0.02	1.01
VFQ*AF	*****			
VSAT*Sz	*****			
VSAT*SS	*****			
VSAT*ISP	*****			
VSAT*AF	*****			

Table 4.18. continued.

TX				
Variables	**Coefficients (B)**	**Wald Statistic**	**p**	**Exp (B)**
Constant	1.23	30.74	<.001	3.41
Sz	*****			
SS	*****			
ISp	*****			
AF	*****			
Column	*****			
Row	*****			
Drop Location	*****			
Trial #	*****			
Age	*****			
SF-12 MCS	*****			
SF-12 PCS	*****			
Dexterity	*****			
VFQ	*****			
VSAT	*****			
VFQ*Sz	2.84E-03	4.24	0.04	1.00
VFQ*SS	*****			
VFQ*ISP	*****			
VFQ*AF	*****			
VSAT*Sz	*****			
VSAT*SS	*****			
VSAT*ISP	*****			
VSAT*AF	*****			

Table 4.18 continued.

MDC				
Variables	Coefficients (B)	Wald Statistic	p	Exp (B)
Constant	8.20	39.92	<.001	3625.87
Sz	-0.85	5.83	0.02	0.43
SS	-1.40	15.33	<.001	0.25
ISp	*****			
AF	*****			
Column	1.68	8.82	<.001	5.37
Row	*****			
Drop Location	*****			
Trial #	*****			
Age	*****			
SF-12 MCS	*****			
SF-12 PCS	*****			
Dexterity	*****			
VFQ	*****			
VSAT	*****			
VFQ*Sz	*****			
VFQ*SS	*****			
VFQ*ISP	*****			
VFQ*AF	*****			
VSAT*Sz	*****			
VSAT*SS	*****			
VSAT*ISP	*****			
VSAT*AF	*****			

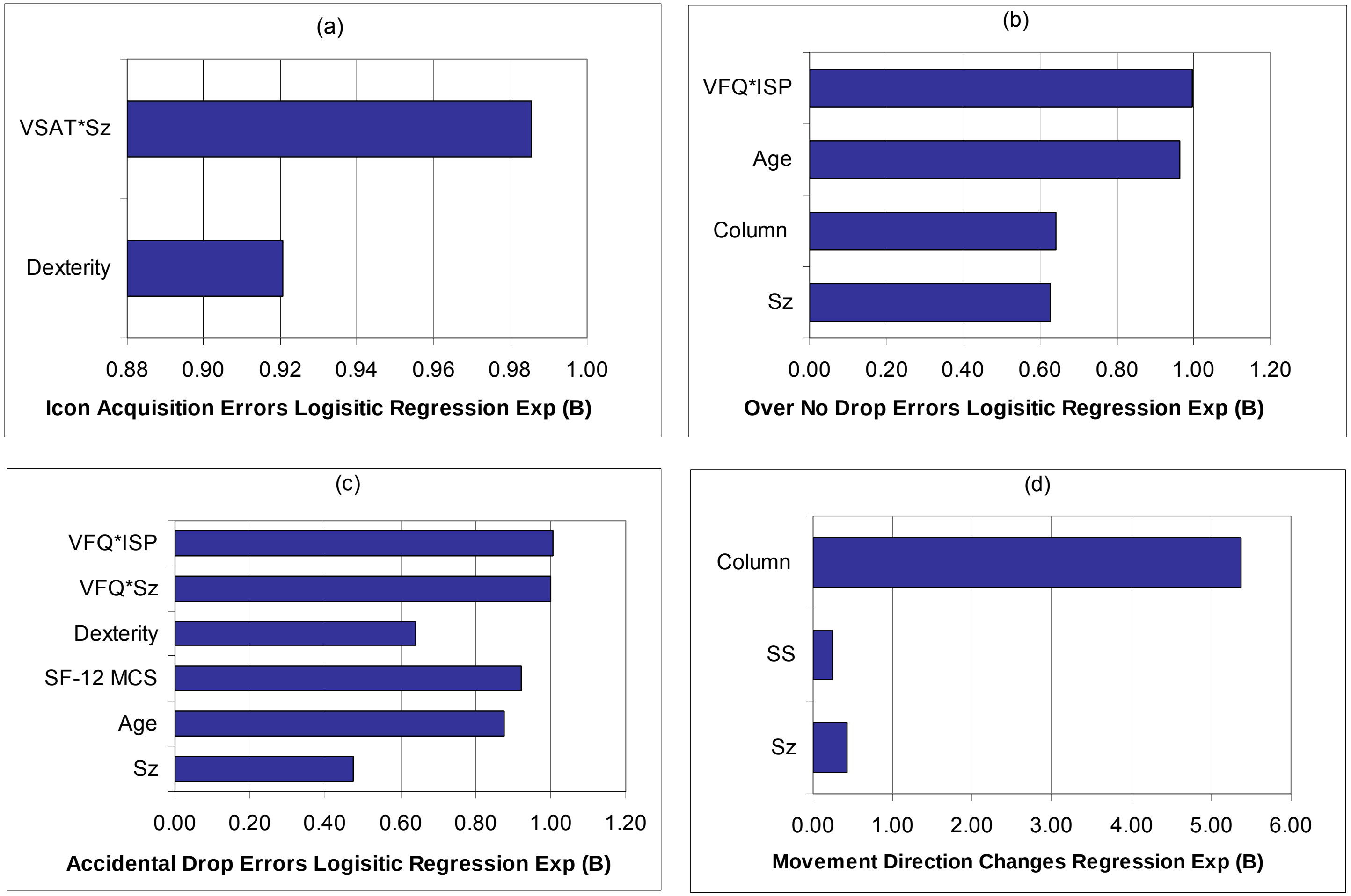


Figure 4.11. Illustration of the relative impact of the predictor variables (non-clinically-acquired) on the likelihood of each error occurring at least once.

Accuracy Measure Outcome Summary

Clinically-acquired Ocular Factors

Outcome #1 Clinically-acquired Predictors

Near Visual Acuity and AMD Severity. In the models of accuracy, there were few cases where clinically-acquired ocular factors were included predictors. However, the impact of NVA was exceptionally powerful in the prediction of AD, and the influence of AMD Score was observed to strongly influence the likelihood for IA. For the model of IA, the likelihood for incidence of this type of error was observed to increase as the severity of the disease worsened. NVA had a similar effect on AD, where diminished NVA increased the likelihood for accidental drops on the approach to the card pile with the icon. It is surprising that none of the ocular factors, even in interaction terms, demonstrated measurable effects on OND or MDC. The models for OND and MDC were largely informed by facets of the interface only.

Outcome #2 Controlled Interface Components

Icon Size. Sz emerged as a persistent influence in the models of accuracy on the desktop. The affect of increased Sz was shown to have positive influence on task accuracy as reflected in the models of IA, OND, AD, and MDC. This demonstrates that there were positive effects of size across all participants for the accuracy on all phases of the task – icon selection, movement and its release.

Icon Spacing. ISp was included as a predictor in the model for OND errors. This model implicates increased ISp to reduce the tendency to commit OND in the release of the icon into the drop pile. Similar to its effect in the handheld task, increased inter-icon spacing had a positive influence over the drop portion of this task and a negative effect on the timeliness with which the task was executed.

Set size. The likelihood for MDC was shown to decrease in the model, based on increased SS. The only other instance of SS influencing accuracy of the task was in the model of TX, in which its interaction with AMD Score also decreased the probability for TX. This implies that SS, influences accuracy of the movement of the icon more than icon selection or release. In addition, increased SS detracted from the efficiency, but was linked to improvements in accuracy, suggesting a speed accuracy tradeoff triggered by the presence of additional distractions.

Auditory Feedback. Perhaps the most compelling outcome of the accuracy analysis is the absence of AF in any of the clinical models. While this does not rule out AF being helpful on a case by case basis, it signals that in the desktop, the AF is not constructive in improving the accuracy on the drop portion of the task, even for those who experienced measurable levels of visual dysfunction. The magnitude of influence of AMD score on icon acquisition, suggests that the feedback may be more beneficial if it had addressed that component of the interaction, rather than the release of the icon. This outcome is contrasted with the bearing AF demonstrated on the efficiency measures of the task. Individuals with AMD were modeled to experience faster TTHT and MT in the presence of AF; the relative performance gains greater at more severe stages of the disease.

Outcome #3 Interface and Visual Factor Interactions

*AMD*Sz.* While main effects were observed for Sz across all participants to influence a lower probability or AD, the effect was amplified as AMD level increased in severity. In other words, as the AMD*Sz term increased in value the likelihood for AD errors diminished. This implies that the increase in size assisted in diminishing the number of premature releases the icon, and suggests that it was easier for the participants to discern the point in time when they were in correct position for an accurate drop. It is surprising that Sz emerged repeatedly as influential on the different measures of task accuracy and that AF did not, particularly in those measures reflecting the accuracy in the ‘drop’ component of the task.

*AMD*SS.* As mentioned previously, the only factor to be included in the prediction of TX likelihood was the AMD*SS interaction term. The likelihood for TX decreased as SS increased with the presence of AMD at increasing levels of severity. It is curious that this was the *only* factor to emerge as a predictor in the model of TX.

Outcome #4 Personal Factors

Age. Age emerged as having a small impact on the models of both OND and AD. Both measures of the success of the release of the icon into the drop pile. Increased age in both models generated a slight decrease in the probability for the errors. Because age was observed in the non-clinical model to inform an increase TT, this result suggests that the older participants exhibited a speed accuracy tradeoff in their performance.

Dexterity. Dexterity only had influence in the model of AD. Increases in dexterity led to a decreased likelihood for AD. This is appropriately reflective of the amount of manual control necessary for the coordination of mouse input with the events on the visual display. Still, in the AD model, the visual factor greatly overrides the impact of dexterity for this population.

Mental Health. SF-12 MCS had a minor influence in the model of IA. The model suggests that the heightened levels of MCS decreased the likelihood for AMD errors during the course of a trial. MCS also emerged as a significant influence on decreased TT, VST, and MT and TTHT. It is likely that the participants were using the mouse pointer during their visual pursuit of the target icon, and missed opportunities in IA in addition to demonstrating difficulty in the selection of the icon, represents times which the icon was not attended to by the participant during their scan of the interface.

Outcome #5 Other interface features

Learning. The impact of trials later in the experimental task did not exhibit the widespread influence on accuracy that was observed on the efficiency measures. That is, only OND included Trial# as a predictor, which held a slight influence in decreased OND errors.

Column. The likelihood for MDC was largely directed by the location of the target icon on the display. Target icons situated in columns further to the right were more prone to a commit MDC errors. The magnitude of this influence was quite strong as compared with the influence of SS and Sz on decreasing MDC probability. In other words, increased SS and Sz did not counteract the negative influence of card icons located in the column further from the drop piles. The location of icons was also included as a predictor of OND likelihood. Icons positioned in the columns further from the drop pile demonstrated a decreased likelihood for OND. This can be interpreted to suggest that icons closer to the drop piles are more prone for OND. This increased difficulty may be due to the difference in the motor movements needed to reach the target: fine and precise movements with the mouse to move from those columns immediately to the right of the drop piles, and large gross movement for those icons further to the right of the display (Darling, Cooke and Brown 1989; Smith, Sharit et al. 1999).

Non-Clinically-Acquired Ocular Factors

Outcome#1: Differences between Clinical and Non-Clinical Models

Table 4.19 presents the differences and similarities between the predictor terms included in the desktop versus clinical logistic regression models. The models were consistent in the prediction of MDC, incorporating Column, SS, and Sz – only interface features. In addition, both models for TX only incorporated one term, which consisted of an interaction between an ocular factor and SS or Sz. The influence of Sz on accuracy was observed in both non-clinical and clinical models, as were the effects of age on OND and AD. Finally, main effects and interaction terms with AF were not included in any of the accuracy models, clinical or non-clinical models.

Table 4.19. Differences between clinical and non-clinical models derived for the accuracy measures.

Model	Consistent	Unique to Clinical Models	Unique to Non-Clinical Models
p(IA≥ 1)		AMD Score SF-12 MCS Sz	VSAT*Sz Dexterity
p(OND ≥ 1)	Age Column Sz	Trial# ISp Sz	VFQ*ISp
p(AD ≥ 1)	Dexterity Age Sz	AND*Sz NVA SF-12 MCS	VFQ*ISp VFQ*Sz
p(TX ≥ 1)		VFQ*SS	VFQ*Sz
p(MDC > 1)	Column SS Sz		

Outcome #3 Controlled Interface Factors

Set Size and Icon Size. The influence of SS and Sz were consistent with their impact in the clinically acquired models. That is, increases in SS or Sz generated a decreased likelihood for OND, AD and MDC. In addition, several interactions between ISp, Sz and the visual factors (VFQ and VSAT) were included as predictors in the model, and help to further explain the observed effects.

Outcome #4 Interactions between Visual and Interface Factors

*VSAT * Sz.* The only accuracy measure to include VSAT as a predictor was IA, which predicted a lower probability for IA in when the interaction VSAT*Sz increased. This suggests that the positive effects of larger icon sizes are more influential for individuals with higher levels of visual attention. Overall, the diagnostic capability of VSAT for the assessment of task accuracy was low on the desktop computer.

*VFQ*ISp.* The interaction term VFQ*ISp was included as a predictor in the models of OND and AD likelihood. In both regressions, increases to VFQ*ISp imposed a slight influence on increasing the probability for the errors. This implies that increases in spacing, for the drop portion of the task, may not benefit those with higher levels of perceived vision. The size of these effects, relative to the influence of other factors such as column and dexterity are quite small.

*VFQ*Sz.* VFQ*Sz was an included predictor in both AD and TX models. It was the only predictor to factor into the likelihood of TX. As the interaction term increased, the likelihood for TX also increased proportionally. This suggests that size of icons can pose challenges, especially for more normally-sighted users in its manipulation around the display.

The effect of VFQ*Sz on AD counteracts the effects observed for Sz. That is, increased value in the interaction term VFQ*Sz led to an increased likelihood for AD, while there was a stronger, overall main effect for Sz to decrease the probability for AD across all users. This suggests that the size of the icon enables a more easily executed drop of the icon into the pile, but that the extent of impact of icon size depends on the visual dysfunction of the user. Those with more pronounced levels of visual dysfunction will benefit more from increased icon size, particularly on the effective release of the icon at its target destination.

Outcome #5 Personal Factors

The influence of both age and dexterity were consistent with the clinical models for accuracy. That is, increased age led to lower likelihoods for AD and OND, and improvements to manual dexterity led to a lower likelihood for AD. An additional effect was observed for dexterity on the model for IA – higher levels of dexterity generated a decreased likelihood for IA. This is logical as the acquisition and release of the icons mandate a high degree of timely visual-motor coordination.

Outcome #6 Other Interface Features

While the influence of Trial# was not observed as an included predictor the non-clinical models, the impact of column was consistent. Target icons placed in the columns further to the right on the display influenced a much higher likelihood for MDC, and a measurable decrease in the probability for OND errors.

4.4 Information Processing

General Summary

Once processed and cleaned, the eye data was evaluated in light of the experimental notes and integrity of the values. It was determined acceptable eye movement and pupil data were gathered from nine of the participants in the desktop PC experimental task. In four cases it was not possible to capture an adequate image of the eye in the time given for the task, or the quality of the tracking was questionable. Of the nine useful sets, the pupillary change was excluded for one participant. In this participants' trial, it was noted the eye tracker had to be adjusted several times through the task, as the participant inadvertently bumped it out of place. This adjustment was noted in the data stream, as pupillary change was observed at an unrealistic magnitude. The adjustments made to refocus the camera after the movement likely placed the camera at a slightly different distance from the eye, generating a different level of magnification. This does not, however, compromise the extraction of the fixation and saccade duration, which are independent of actual scan patterns.

Table 4.20 summarizes the distributions for the eye movement summary measures, fixation duration, saccade duration, and saccade to fixation duration ratio. The percentage of actual fixations recorded by the eye tracking system deviated between participants, based on the quality of the eye image capture, aspects of their eye, or the overall time on the task. Because of this, the measures reported are not based on the relative quantity of fixations, but on summaries of the duration of saccades and fixations.

The relative distributions of the saccade and fixation duration for all participants are shown in Figure 4.12. This graphic demonstrates, as in the handheld task, that saccade duration was prone to a higher degree of variability across the participants. Those ratios at levels above one are indicative of the eye movements, which were spent largely in the pursuit of information to fixate on- a sign of increased breadth versus depth of search. This is also an indication of inefficiency in the extraction of information from the display. Figure 4.13 illustrates saccade to fixation ratio term across the participants.

Table 4.20. Summary of physiological measures of information processing.

Physiological Measures	
Fixation Duration (seconds) (n = 9)	Mean (Std. Error): .41 (.025) Median: .42 Minimum: .25 Maximum: .512
Saccade Duration (seconds) (n = 9)	Mean (Std. Error): .484 (.357) Median:132 Minimum: <.051 Maximum: 3.33
Saccade to Fixation Ratio (n = 9)	Mean (Std. Error): 1.74 (1.43) Median: .314 Minimum: .12 Maximum: 13.20
Pupillary Response (mm) (n=8)	Mean (Std. Error): 2.06 (.258) Median: 2.146 Minimum: .65 Maximum: 2.96

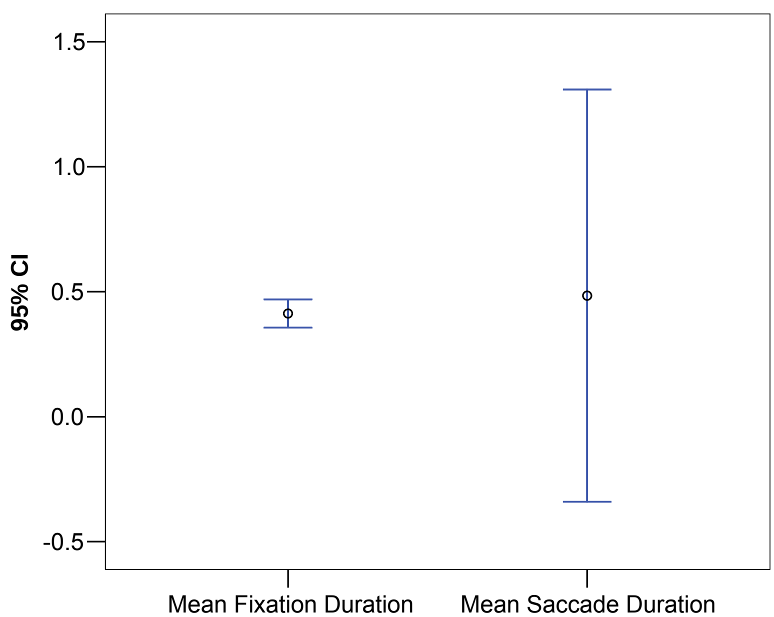


Figure 4.12. Distribution of fixation duration and saccade duration for participants.

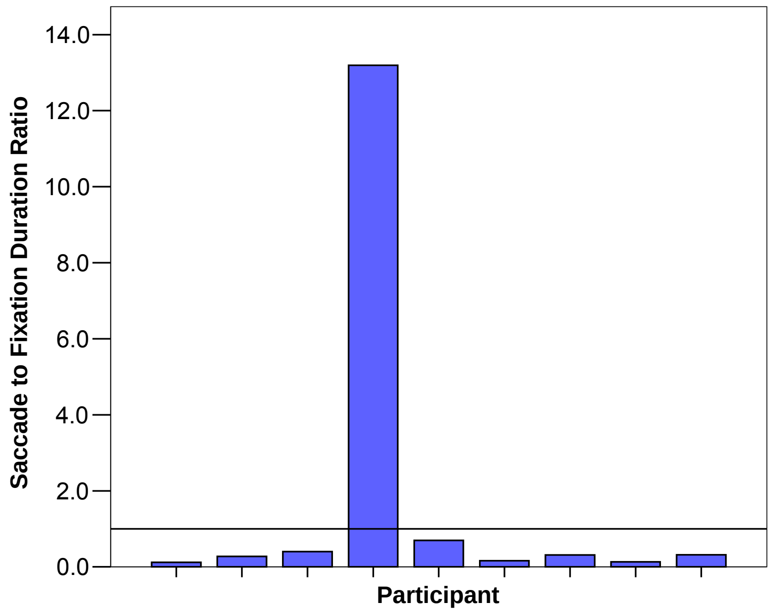


Figure 4.13. Saccade to fixation duration ratio across participants.

The application of the eye tracking system for the analyses of the desktop computer interaction enabled the synchronization between the eye data stream and with events in the interface. This was not possible with the handheld computer experimental setup, due to technical limitations. For the desktop, this enabled the consideration of the impact of the presence or absence of auditory feedback on eye movement and pupillary change. Figure 4.14 illustrates a summary of fixation and saccade duration, along with the saccade the summary statistics between the two conditions on the three variables.

Pupillary response provided an indication of the mental workload encountered by the participants during the task. A time based measure, pupillary response, as calculated by Backs and Walrath (Backs and Walrath 1992), was used, and reported in mm change. Figure 4.15 illustrates the change in pupil diameter, or pupillary response, per participant with the presence and absence of auditory feedback.

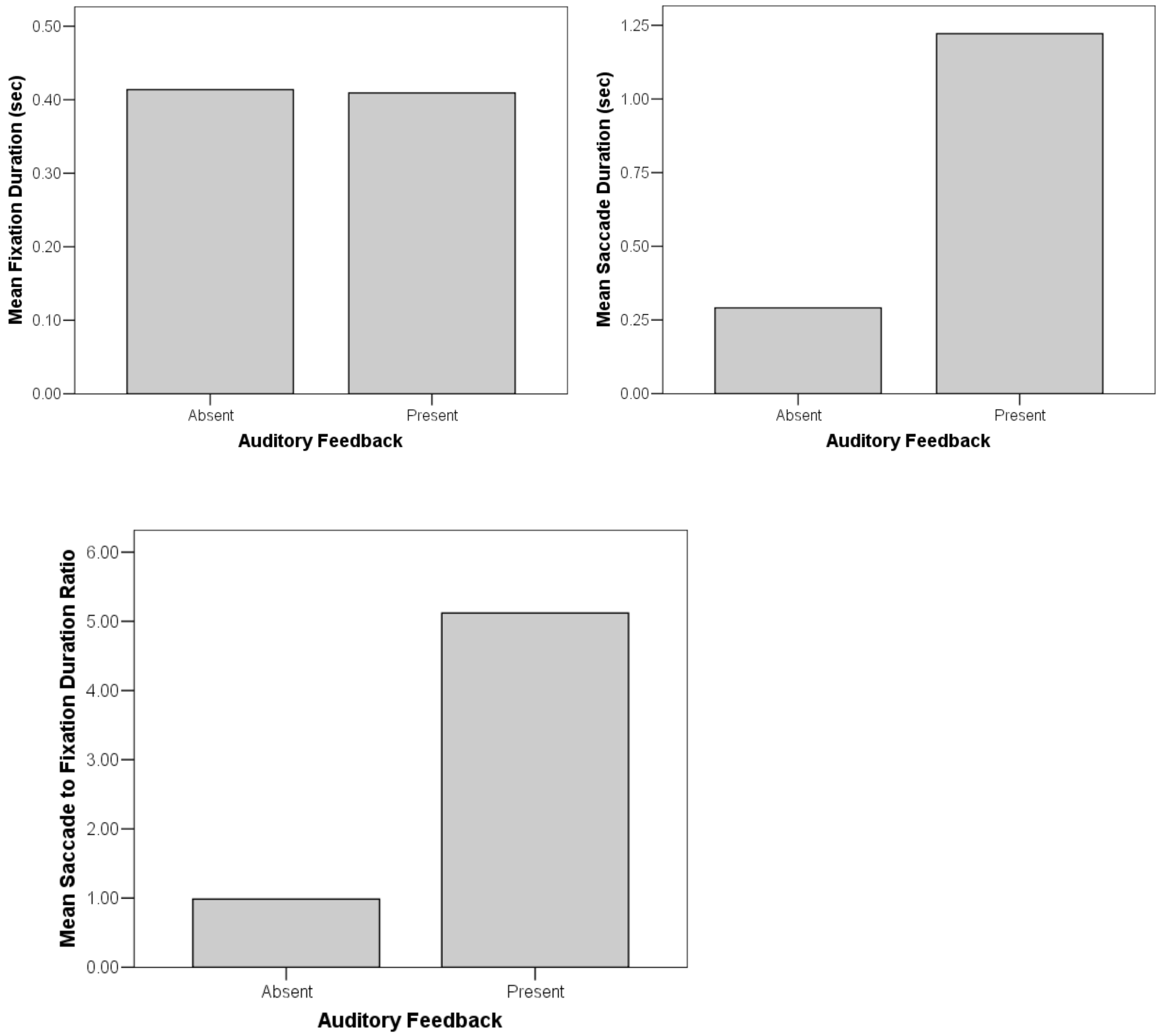


Figure 4.14. Comparison of fixation and saccade duration between the condition with auditory feedback present and absent.

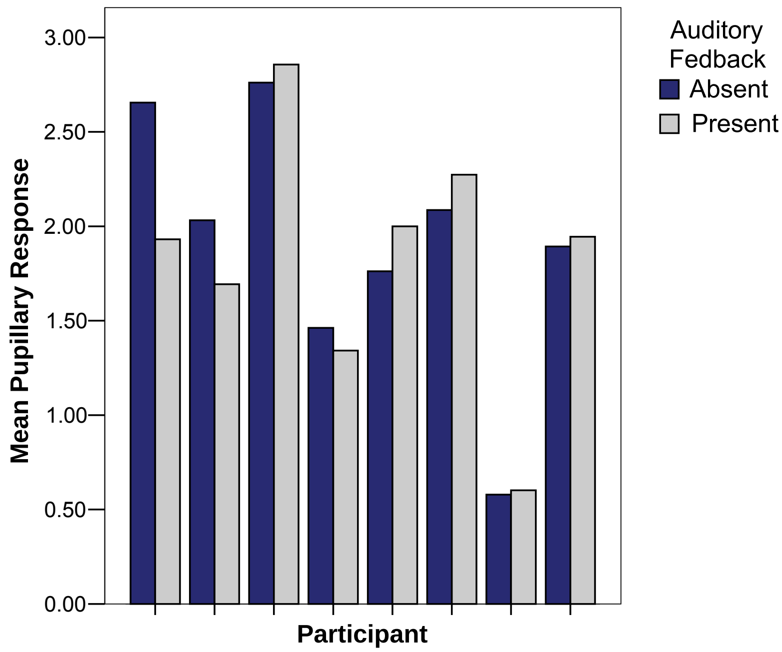


Figure 4.15. Illustration of the pupillary change for each participant with and without auditory feedback.

Analyses: Information Processing

The eye movement and pupillary response data, unlike the other dependent measures, were not captured at all levels of the independent interface factors (SS, ISp and Sz) and only between AF conditions, present or absent. This, combined with the limited number of acceptable cases remaining after data cleaning, indicated that the data set was not well-suited for regression analysis. In addition, the nature of the data, and the small sample size, did not meet the assumptions of parametric statistics, even after attempts at the transformation of the dependent variable.

Instead of regression models, non-parametric statistical comparisons were applied to the eye data, using the Mann-Whitney test (Field 2000; Pallant 2003) . Two groups were created for each eye tracking metric, based on the midpoint (50^{th}

percentile) of the data from the 9 participants (8 participants in the case of the pupillary change metric). While a somewhat unconventional approach, this technique is commonly used in psychosocial research with natural experimental designs such as questionnaires (e.g., Miller 1987). These groups are summarized in Table 4.21 and Figure 4.16. Mann-Whitney comparisons showed that the two groups differed significantly on eye tracking measure from which they were derived. (e.g., Group 1 for Mean Fixation Duration had significantly lower Mean Fixation Duration than Group 2), at an alpha level of .05. For each measure, comparisons between the two groups were made on all the clinically-based ocular measures, non-clinically-acquired ocular factors and participant-based factors.

Table 4.21. Summary of Eye movement metric group classification.

	Fixation Duration (Mean)	**Saccade Duration (Mean)**	**Saccade to Fixation Ratio**	**Pupillary Change (mm)**
50th percentile	0.421	0.132	0.314	2.15
Group 1	< .421 n = 5	< .132 n = 5	<.314 n = 4	≤ 2.15 n = 4
Group 2	≥ .421 n = 4	≥ .132 n = 4	≥.314 n = 5	> 2.15 n = 5
Mann-Whitney Comparison	Z = -2.45 p = .014	Z = -2.45 p = .014	Z = -2.45 p = .014	Z = -2.31 p = .021

< or > denotes the measure for that group as being less than or greater than the median.

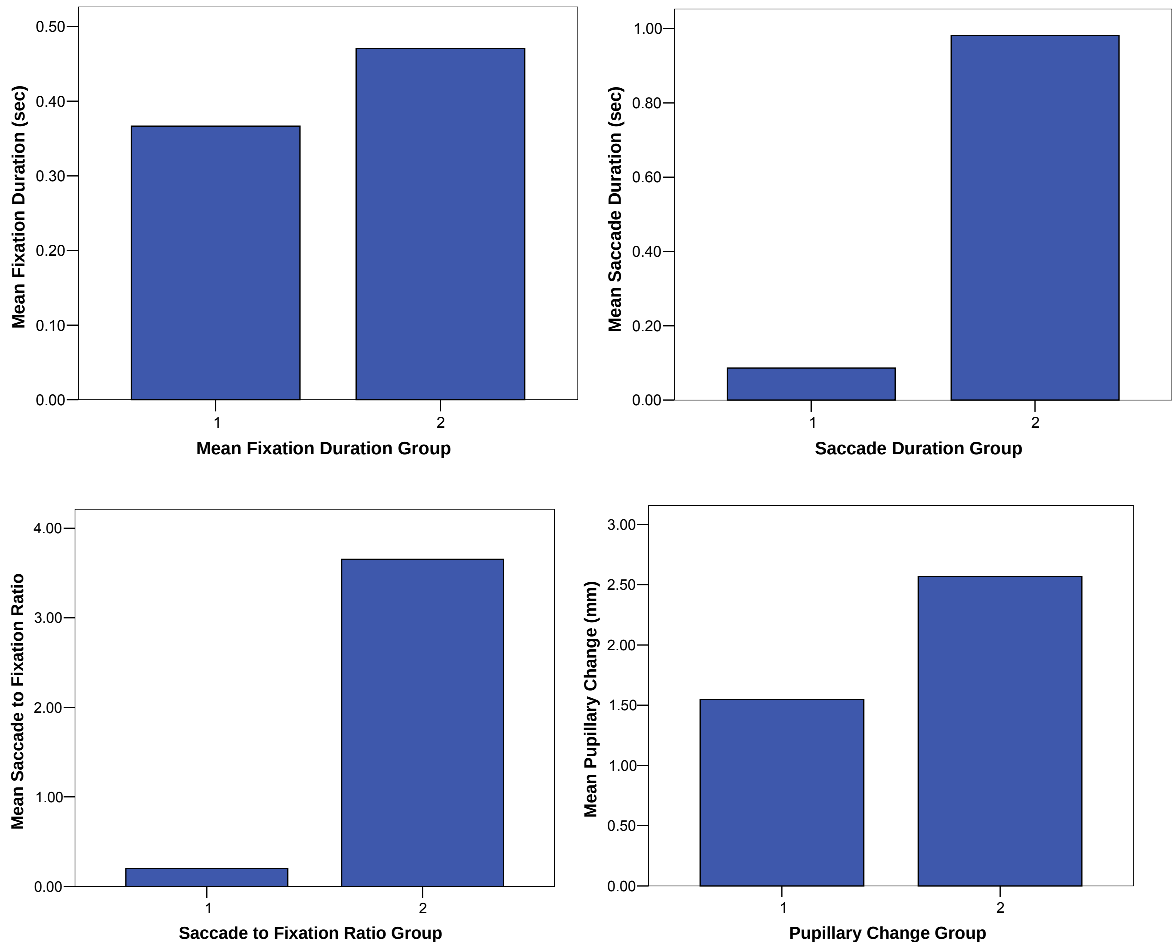


Figure 4.16. Summary of groups derived from eye tracking measures.

Results: Information Processing

As stated, Mann-Whitney tests were used to compare the differences in personal factors, clinically-acquired factors, and non-clinically-acquired factors between the two groups on each eye tracking metric. Secondly, the eye movement and pupillary response data was analyzed between the auditory present and absent conditions using the Wilcoxin Signed Rank test, for comparing two related samples.

The comparisons between the various information processing groups on personal and ocular factors did not result in any significant differences (alpha level = .05). This was surprising, as the same analyses for the handheld component yielded a handful of interesting outcomes which offered additional insight into the underlying mechanisms driving interactions.

The eye stream data was then considered for effects attributable to the presence of AF. The Wilcoxin Signed Rank test was used, across all participants to determine if there were differences the underlying information processing that occurs in the presence versus the absence of auditory feedback (Field 2000). Mean saccade duration ($Z= -1.970$, $p = .049$) emerged as significantly different between the trials versus those without AF, the difference illustrated in Figure 4.14. According to the result, mean saccade duration was significantly longer in the presence of auditory feedback across all participants. This is a particularly compelling result, as it clarifies the effects of AF in the efficiency metrics, which suggested the propensity for longer visual search in the presence of visual feedback but improved MT and highlight times.

Further analyses were conducted to determine if the differences in saccade duration emerged based on the severity of ocular dysfunction. Figure 4.17 illustrates the mean saccade duration, categorized into three groups based on the severity of AMD; 1) healthy controls (AMD Score = 0); 2) Mild AMD (AMD Score .25-1); 3) Severe AMD (AMD Score greater than 1). While statistical tests did not reveal significant differences within these groups on the duration of saccades, the graphic suggests that the controls were more prone to longer search times in the presence of AF. From the participants with auditory feedback, just two were visually-healthy controls, four were classified with between .25-1.00 AMD Score, and three were classified with 3.00 or 4.00 AMD Score. It is anticipated that the inclusion of more controls would likely amplify the effects.

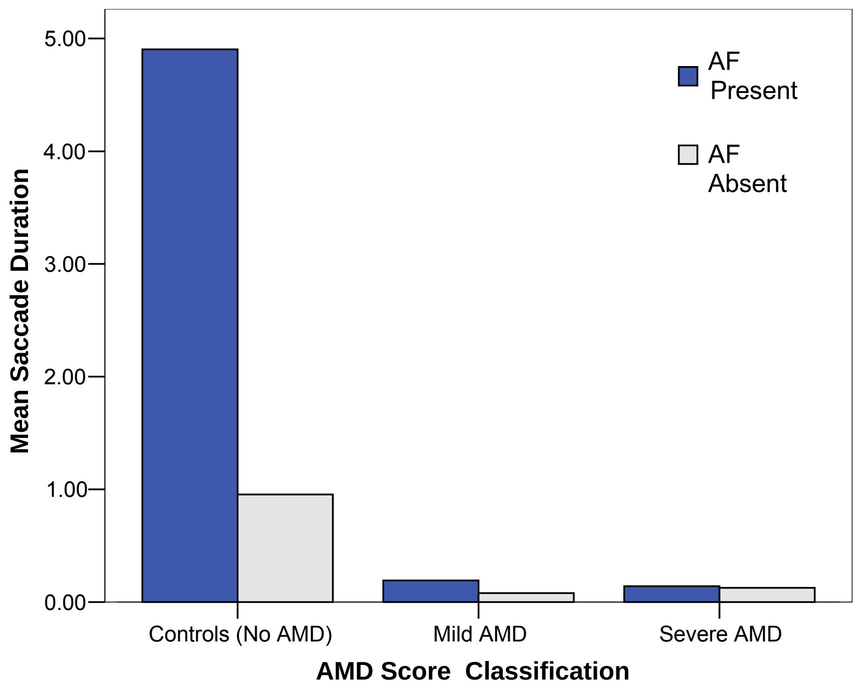


Figure 4.17. Mean saccade duration between AF conditions based on AMD Score.

Information Processing Outcomes

Outcome: Auditory Feedback and Information Processing

The analyses of information processing facilitated by the eye movement measures implied that visual search was less efficient in the presence of auditory feedback, with greater time spent between fixations. This result, while not significant within AMD Score groups, showed to be particularly salient for those participants in the control group. This lends additional explanation for the ‘distraction’ effect observed for AF in the analyses of the VST. The participants were less direct in their search for icons in the presence of AF.

4.5 Consolidation of Model

The considerable number of results presented in the analyses of task efficiency, accuracy, information processing, for both clinically and non-clinically-acquired visual factors, personal factors, and task factors generated a large number of interesting results. However, their global interpretation is needed to effectively glean the most compelling contributions to the existing knowledge base. Tables 4.22-4.25 provide the opportunity to evaluate the comprehensive set of results, and draw out the most significant patterns.

A separate table summarizes each class of predictor variables and the interaction terms. The results within each are organized by the various dependent variables that were captured. When appropriate, the source of the result is noted; whether it was generated from the clinical models or the non-clinical models. For each predictor, the tables demonstrate the general relationship with the outcome measures based on increases in the predictor value. Lastly, the table includes an indication of the directional effect observed on the outcome measures in the presence of increases to the predictor terms. These tables provide a useful way to extract the general trends emergent from theses analyses, while the bar graphs in each section are more useful in quantifying the magnitude of the impact on the outcome measures.

Table 4.22. Summary of outcomes attributed to the visual factors.

Clinical	**Predictor: AMD Score** Increased disease severity level (worse)	
	Increasing	IA
	Decreasing	MT
	Predictor: CS Score Higher, improved contrast sensitivity	
	Increasing	FTHT
	Predictor: NVA Score Decrease in visual acuity, worse vision	
	Increasing	TT
		VST
		MT
		TTHT
		FTHT
		AD
Non-Clinical	**Predictor: VFQ Overall** Improved perception of visual function	
	Decreasing	VST
	Predictor: VSAT Percentile Higher percentile for visual attention (improved)	
	None	

Table 4.23. Summary of outcomes based on the independent task-related factors; the source of the outcome, clinical or non-clinical models, is identified with a respective checkmark in the appropriate column.

Predictor: AF AF absent - AF present		**Clinical**	**Non-Clinical**
Increasing	VST	✓	
Decreasing	TTHT		✓
Predictor: SS Increasing the number of icons		**Clinical**	**Non-Clinical**
Increasing	TT	✓	✓
	VST	✓	✓
Decreasing	MDC	✓	✓
Predictor: ISp Increasing the inter-icon spacing		**Clinical**	**Non-Clinical**
Increasing	TT	✓	✓
	VST	✓	✓
	MT		✓
Decreasing	OND		✓
Predictor: Sz Increasing the icon size		**Clinical**	**Non-Clinical**
Increasing	MT	✓	
Decreasing	VST	✓	✓
	FTHT		✓
	IA	✓	
	OND	✓	✓
	AD	✓	✓
	MDC	✓	✓

Table 4.24. Summary of the interactions between the visual factors and the interface independent variables (sumamrized for increases in the value of the predictor variable).

Group	Predictor	Direction	Variable
Clinical	**Predictor: AF*AMD Score**		
		Decreasing	VST
			MT
			TTHT
	Predictor: SS*AMD Score		
		Decreasing	TX
	Predictor: ISp*AMD Score		
		None	
	Predictor: Sz*AMD Score		
		Decreasing	TT
			VST
			AD
Non-clinical	**Predictor: AF*VFQ Overall**		
		None	
	Predictor: SS*VFQ Overall		
		None	
	Predictor: ISp*VFQ Overall		
		Increasing	OND
			AD
	Predictor: Sz*VFQ Overall		
		Increasing	VST
			AD
			TX
		Decreasing	TT
	Predictor: AF*VSAT Percentile		
		Increasing	TTHT
	Predictor: SS*VSAT Percentile		
		None	
	Predictor: ISp*VSAT Percentile		
		None	
	Predictor: Sz*VSAT Percentile		
		Increasing	MT
			TT
		Decreasing	IA

Table 4.25. Summary of the impact of personal factors and their impact on the dependent variables.

PERSONAL FACTORS			
Predictor: Age Older age		**Clinical**	**Non-Clinical**
Increasing	TT	✓	
	MT	✓	✓
Decreasing	VST		✓
	OND	✓	✓
	AD	✓	✓
Predictor: Dexterity Improved manual dexterity		**Clinical**	**Non-Clinical**
Increasing	TTHT	✓	✓
	FTHT	✓	✓
	MT	✓	✓
Decreasing	TT	✓	✓
	VST	✓	✓
	AD	✓	✓
	IA		✓
Predictor: SF-12 PCS Improved rating of physical health		**Clinical**	**Non-Clinical**
Increasing			
Decreasing	TT	✓	✓
	VST	✓	
	MT		✓
Predictor: SF-12 MCS Improved rating of mental health		**Clinical**	**Non-Clinical**
Decreasing	TT	✓	✓
	VST	✓	✓
	MT	✓	✓
	TTHT	✓	✓
	FTHT	✓	✓
	AD	✓	✓
	IA	✓	

4.6 Exit Survey, Subjective Participant Responses

After completion of the task, participants were asked a series of questions regarding their experience. These included questions concerning their perception of their performance and workload during the task, their comfort with the equipment, and their opinions of the various interface manipulations.

Participants were positive about their experience. The participants rated their comfort level with the task and computer as *very comfortable* (n=4), *comfortable* (n= 6), and *neither comfortable nor uncomfortable* (n= 2). One individual did rate the experience as uncomfortable, but stated their comfort level was a result of the weight of the trial frames on their face.

When asked to rate what they liked best about their experience, the participants were in general very opinionated. Several actually enjoyed the task and using the cards. Table 4.26 provides a summary of participants' response to the questions: *What did you like best about this experience*, and *What did you like least, or dislike about this experience?*

The positive effect of using the playing card icons is observed in the 'likes' category, where participants related the task to playing a game, and made comments about their strategies, or how they liked the challenge. The aversions were, for the most part, related to their discomfort with the trial frame glasses. However, the experimenters continually made adjustments to the frames as needed over the course of the experimental trial to improve upon their fit to ensure patient comfort. Other aversions were classified as task-related, such as the expressed difficulty faced with tracking the mouse pointer on the screen, and the aversion for to the smallest icon size.

Participants also rated their overall performance, their perceived difficulty with the task, their effort in the execution of the task and the frustration experienced. Their response to each category provided an indication of their perceived mental workload, and as mentioned in Chapter 3, and were derived from subscales used in the NASA TLX system (Hart and Staveland 1988). Figure 4.18 illustrates participants' responses to these questions of perceived workload. Each workload factor was rated by the participants on a scale from 0 to 10, low to high, similar to the NASA TLX scoring, but on a smaller relative scale (0-10 instead of 0-100), according to the convention noted on each graphic.

Table 4.26. Summary of participant likes and dislikes about their experience with the handheld PC and the experiment.

Likes	Dislikes
Nothing.	*Nothing. (n=5)*
It was comfortable.	*Smallest cards. It was hard to distinguish between suits; especially spades and clubs.*
It was a very good experience - It was something different to do. I felt good about it.	*Trouble with the arrow when it went on the card I couldn't see it. It was very hard for me to place it.*
It was redundant- easy.	*The small cards.*
Finding the card.	*The glasses were uncomfortable (n=4)*
Fun game.	*Became boring; Went on too long.*
It made (me) aware of what was going on. I had to pay attention, and coordinate vision and listening.	*I didn't like it or dislike it. It was a wonderful tool.*
It was helpful and interesting. I was glad to have the experience.	
Just that I played cards, and I haven't been able to in a long time.	
It was a lot of fun, a learning tool.	
Your (the experimenters') company.	
It was ok.	
That it was finished.	
Not much.	

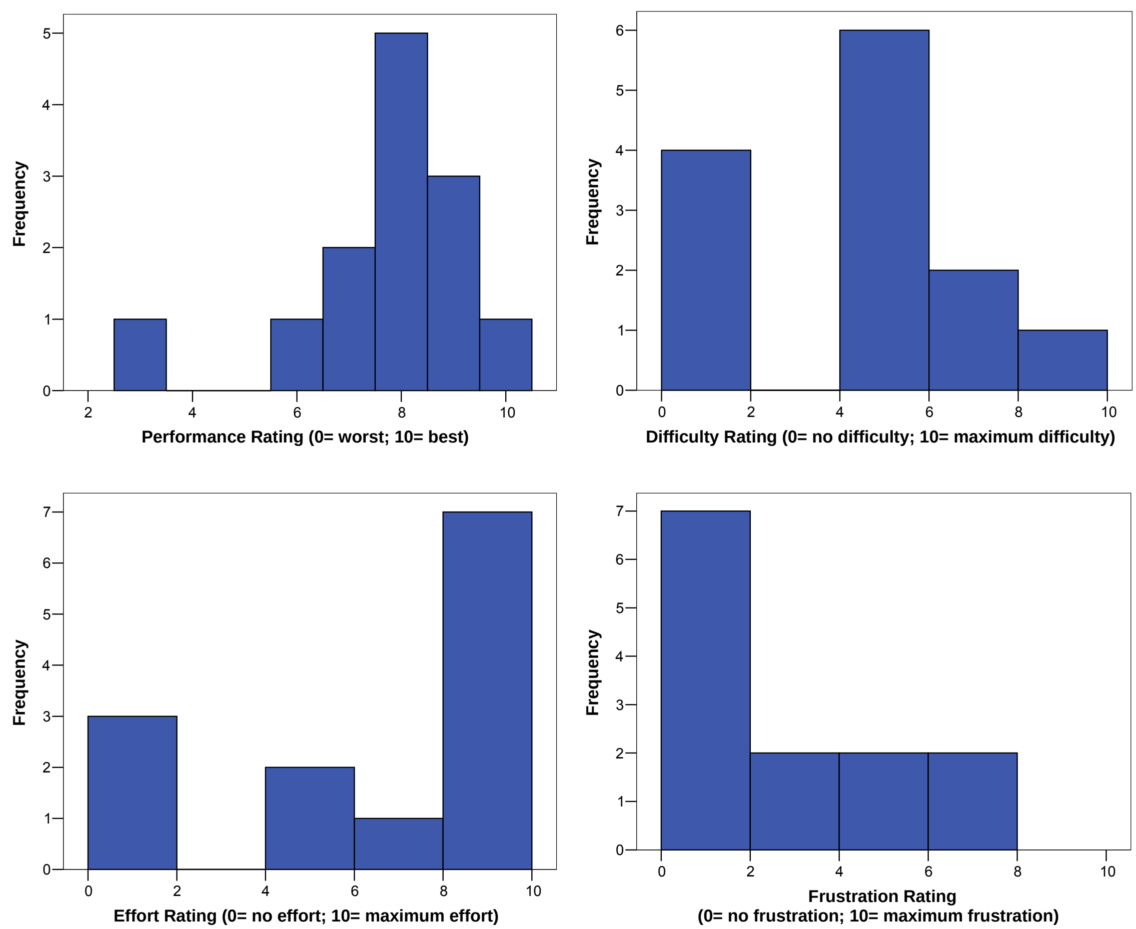


Figure 4.18. Summary of responses to perceived workload subscales.

All but one participant rated their overall performance on the task with a score of 6 or better, with the majority attributing a score of 8 to their performance. This accounts for the low error rate observed in the performance assessment of task accuracy. However, the distributions of the ratings allotted to difficulty and effort are more indicative of the differences observed in task efficiency. More than half of the participants rated the task difficulty above a 5, and the same was true for effort. So, while they felt that they were successful, the challenge or workload encountered to achieve that success was quite varied between participants. Participants overall demonstrated a high success rate in task completion (correct card icon to correct drop pile), but the rate at which they completed the different components of the task, and the occurrence of errors of commission during the task differed, were largely driven by personal, ocular and interface related factors.

As these are perceived rates of mental workload, it is of interest to compare these ratings to the measure of workload and information processing captured through eye movements and pupil diameter. To this end, the existing groupings for mean fixation duration, mean saccade duration, saccade to fixation ratio, and pupillary change were used to detect differences in the perceived task workload subscales. The Mann-Whitney U non-parametric test for between group comparisons was used. None of these tests detected significant differences (alpha= .05) between these groups on any of the four subscales. In addition, non-parametric correlations were performed using Pearson's rho and Kendall's tau, neither of which detected significant differences between any of these groups for the perceived workload responses (alpha= .05). This suggests that even though the subjective measures are appropriate and necessary to consider how the participants felt about their interactions, they are not diagnostic enough to endorse specific interface characteristics.

The participants' reactions to the sound were consistent with its impact on performance. That is, there was a mixture of reactions to it- both positive and negative. In response to an enquiry regarding their comfort with the sound, none of the participants rated the sound as *uncomfortable* or *very uncomfortable*. Participants ratings on comfort with the sound were evenly divided amongst *very* *comfortable* (n= 4), *comfortable* (n= 5), and *neither comfortable nor uncomfortable* (n= 4).

The question: *How helpful was the sound to your completion of the task?* received a similar set of responses. None of the participants rated the sound as *unhelpful* or *very unhelpful/distracting*. The majority of the participants were impartial with regard to sound helpfulness. Eight of the participants claimed the sound was *neither helpful nor unhelpful*. The remaining five participants rated the sounds as very helpful (n= 2) or helpful (n= 3). The participants' perception of their comfort with the AF and their opinion of how helpful it was in accomplishing the goals of the task were consistent with the attitudes of the participants who completed the handheld experiment. The pie graphs shown in Figure 4.19 summarize the participant responses to the questions concerning AF. In addition, the participants provided free responses on their general thoughts on the sound, which appear in Table 4.27, organized by positive, negative and neutral opinions.

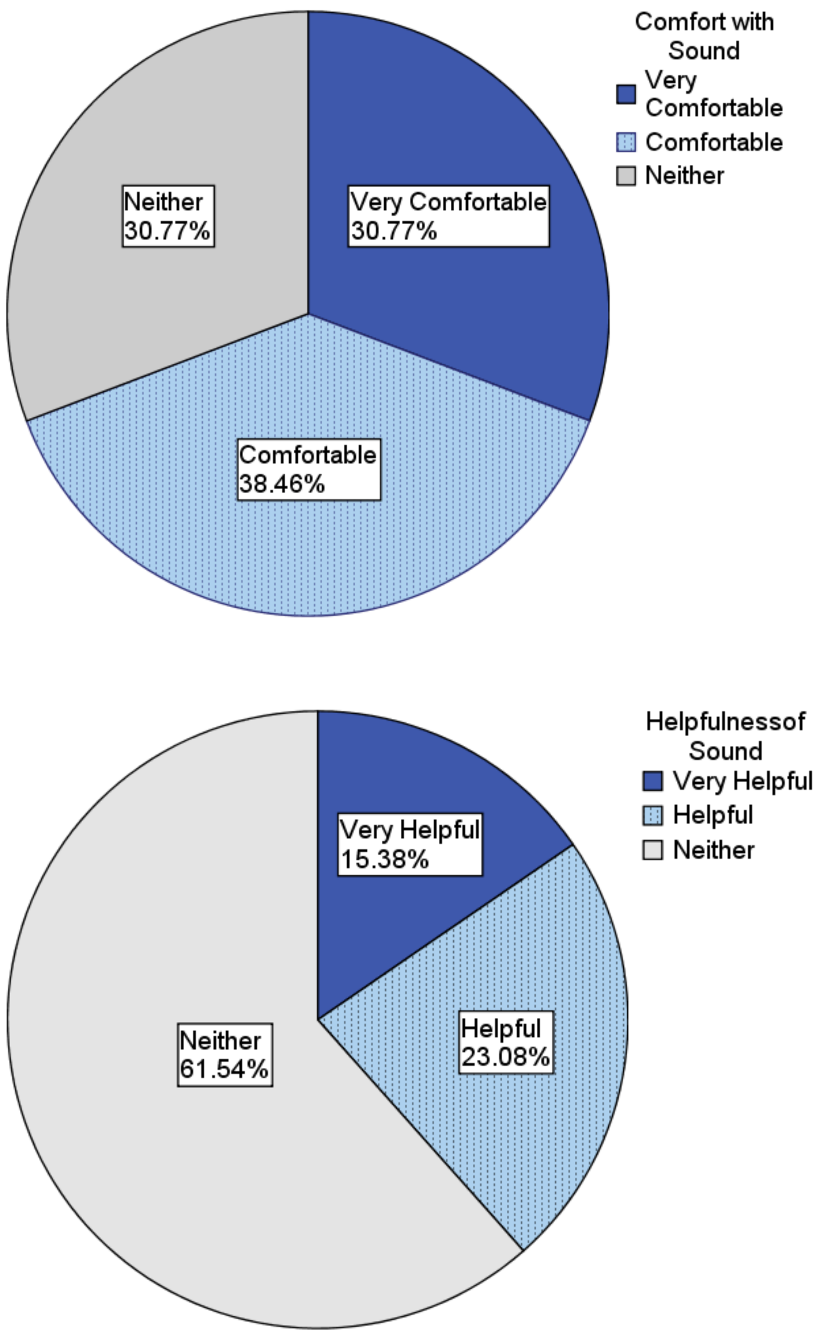


Figure 4.19. Summary of participant response to questions regarding their comfort level with the auditory feedback and their perception of the auditory feedback helpfulness to the task.

Table 4.27. Participant opinions of auditory feedback, verbalized responses.

	Participant Opinions of the Auditory Feedback
Positive	*Good, helpful.*
	It could help you.
	It was all right. At least you knew you had the card in the right place.
	It was good.
	It was fine.
Negative	*Good sound - sounded like cards. Didn't think it was very helpful.*
	Irrelevant, really.
	I didn't feel that the sound added great benefit.
	Gave you a clue about placement, not important.
	Sound didn't mean a thing; Had to concentrate harder on visual.
	Just a zero attention getter; Didn't release based on sound my actions were based on what my eyes saw on the screen.
Neutral	*Didn't bother me - when I was concentrating, you don't really notice it.*
	I didn't pay attention to it.

Further analyses were run on these perceptions of the auditory feedback, to determine any similarities between perception of the feedback and the likelihood of the feedback influencing performance. Based on their response to the question concerning the helpfulness of the auditory feedback, participants were assigned to two groups. Group 1 consisted of those participants who responded as very helpful or helpful; Group 2 was comprised of participants who responded with *Neither*, *Unhelpful*, or *Very Unhelpful*. Comparisons were made between these two groups using the Mann-Whitney test for non-parametric comparisons. The comparisons made in consideration of VSAT, VFQ Overall, and AMD Score did not reveal significant differences between the groups on those visual factors (alpha = .05). In addition, a chi-squared test comparing controls and AMD with the two groups did not reveal significant differences based on that factor.

Participants were also asked to provide their opinions on their preferred icon size. The distribution of preferences for icon size is illustrated in the pie graph in Figure 4.20. None of the participants asserted a preference for the small icon size, the one that mimicked the handheld icon size. The majority of the participants preferred the largest icon size, which was larger than the typical windows icon.

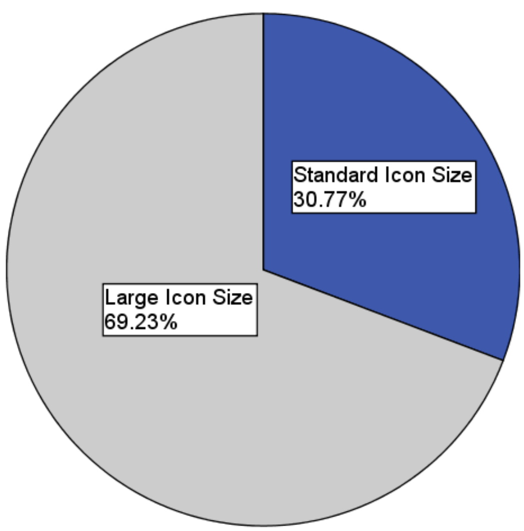


Figure 4.20. Participant icon size preference.

Analyses were performed to explore differences between the participants based on their icon size preference. Mann-Whitney non-parametric statistical comparisons were used to compare the groups based on ocular and personal factors. The only factor to emerge as significantly different between the icon size preference groups was Overall VFQ score ($Z=-2.16$, $p = .031$), as shown in Figure 4.21. This result suggests that the group of participants whose preference is for the standard Windows icon size had a significantly higher Overall VFQ Score, or higher level of perceived visual function with respect to their daily activities. Interestingly, the interaction of VFQ*Sz was included as a significant predictor in the regression models predicting efficiency, and also AD and IA.

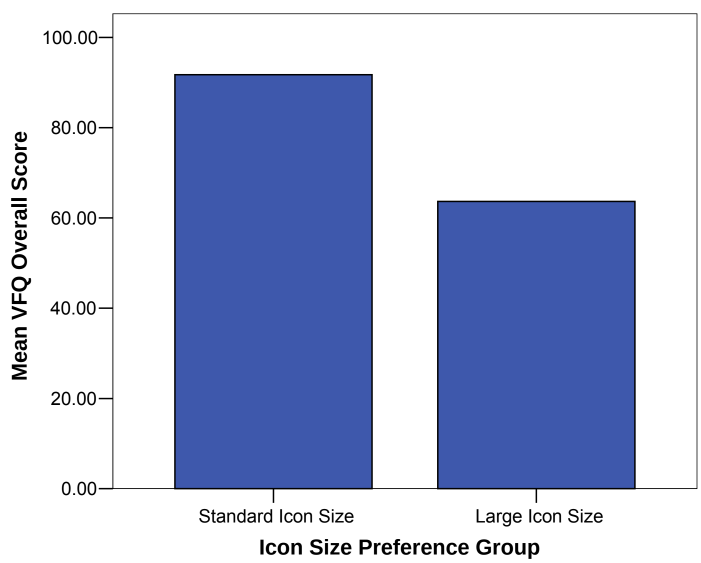


Figure 4.21. Mean VFQ score according to icon size preference size.

4.7 <u>Conclusions</u>

The conclusions for the desktop PC interactions are presented in the following section, to answer the hypotheses established in Chapters 1 and 2. Table 4.28 presents the outcome related to each hypothesis, and the supporting evidence that emerged from the analyses. Each table contains a row with cells labeled *Efficiency, Subjective Response, Information Processing, and Subjective Response.* These refer to the class of measures that supported the conclusion. If a measure informed the conclusion, a check mark will appear in the appropriate column.

Table 4.28. Hypothesis summary.

<table>
<tr><td colspan="5"><u>Hypothesis 1</u>: For all users, there is a point of diminishing return for performance gains attributed to increases to icon size, set size and spacing.</td></tr>
<tr><td rowspan="2">Contributing Results</td><td>Efficiency</td><td>Accuracy</td><td>Information Processing</td><td>Subjective Response</td></tr>
<tr><td>✓</td><td>✓</td><td></td><td>✓</td></tr>
<tr><td>Supporting Evidence</td><td colspan="4">*The impact of increased Sz, SS, and ISp was observed to be a main effect across all participants on both models of accuracy and efficiency and consistently so between the clinical and non-clinical models.

*Increases to both SS and ISp contributed to slower VST and TT; they also contributed to a decreased likelihood for MDC and OND (non-clinical model only), respectively.

*ISp was also observed to negatively affect MT for the non-clinical model only. This indicated larger spaces between the icons contributed to a longer time for participants to move the card to the pile.

*Sz demonstrated, for the most part, very positive influence on task efficiency and accuracy, with the exception of MT. Increased icon size influenced faster VST, a lower likelihood for OND, AD, and MDC in clinical and non -clinical models.

*Sz was shown to influence a lower probability for IA and slower MT in the clinical model.

*For the non-clinical model, Sz contributed to a decreased likelihood for FTHT to exceed the 85th %tile for this sample population.

*Several interactions were observed between Sz and the visual factors.

*Participants were more likely to skip trials with the smallest icon size.

*Participants, in the exit surveys reported a preference for the largest and mid-size icons - none of the participants preferred the smallest icon size, which mimicked the icon size from the handheld task.</td></tr>
</table>

Table 4.28. continued.

Conclusions	*For the desktop task, increased ISp or SS impose decrements across all participants in terms of efficiency but improved the accuracy in the manipulation of the icons on the display. The decreased efficiency with increased ISP contradicts the findings of Everett and Byrne (2004), who measured less efficient icon search when inter-icon spacing decreased, observed with a visually healthy, young population. *The only negative effect observed for Sz, across all participants, was its influence on MT in the clinical model. However, because the strength of the model for MT was questionable, the result is taken light *The more important outcome was that there were essentially no trade-offs in the use of large icons observed for this task and participants. However, a point of diminishing return for increased icon size may be observed in tasks which are highly dependent on the quantity of information presented at one time in the visual field. This is likely the case in tasks that command a high degree of information visualization, where the synthesis of several GUI components enables the best comprehension of the system state. *The results suggest, for good design, that ISp and SS should be kept as compact as necessary in visual search task, but that those tasks requiring the manipulation of icons should carefully consider the proximity of the icons and the number presented at one time on the display

Table 4.28. continued.

H1a. The point of diminishing return is dependent on a user’s visual capacity, including clinical measures of vision, visual attention, and subjective visual functioning.

H1b. Icon Size, set size, and spacing will influence the components of the interaction (e.g., visual search, icon acquisition, dragging and dropping) in different ways and magnitudes, also dependent on the visual capacity of the participant.

Contributing Results	**Efficiency**	**Accuracy**	**Information Processing**	**Subjective Response**
	✓	✓		
Supporting Evidence	*The impact of Sz on users' performance in the model was very much dependent on the aptitude of the visual sensory channel, measured both by the clinical and non-clinical ocular factors. However, the impact of SS and ISp on the interaction efficiency and accuracy was less a function of visual aptitude. *The only interactions observed for SS and ISp were: SS*AMD on decreased TX; ISp*VFQ on increased OND and AD. However, ISp*VFQ was the only predictor term included in the TX model.			
	*The interaction between Sz and AMD was shown to decrease TT, VST, and the likelihood for AD, demonstrating influence across different components of the task. Participants with AMD, and specifically those diagnosed with the more severe cases of the disease demonstrated greater performance gains with the larger icons. *The impact of the interaction between AMD and Sz was particularly salient in the model of TT, in which the interaction demonstrated the capacity to counteract the effects of NVA slowing down trial completion. *The interactions for Sz and AMD, for the clinical model, more heavily influenced the efficiency of the task, rooted in improved visual search time. *The models based on the non-clinically-acquired ocular factors, VSAT and VFQ, showed a modest effect for these assessments. The VSAT interaction with Sz suggests that the individuals with higher visual attention scores experience exaggerated effects of Sz slowing TT and MT beyond what those individuals with lower VSAT scores experience. This could also be interpreted to convey that those with lower VSAT scores could realize a performance gain at these larger icon sizes. The absence of main effects for Sz in either of these non-clinical models makes this interpretation limited. *The model of VST was more telling. Those with higher perceived levels of VFQ were modeled to terminate visual search at a faster rate, and all participants were modeled to find the target icon more quickly in the presence of larger icons. However, as the interaction term VFQ*Sz increased in value, the model reflects a slower rate of search completion. This strongly suggests that those individuals with improved VFQ scores do not benefit as much from Sz increases as those with lowered perceptions of visual function. Importantly the extent of the related influence does entirely negate performance gains from Sz.			

Table 4.28. continued.

Conclusions	*Increased icon size and magnification can be helpful, particularly for more severe levels of visual impairment and ocular disease. The context of the task considered in this experiment, Sz did not most part did not impose performance decrements across all participants, with one exception (the influence of VSAT on VST). Still, the application of this solution should be done judiciously, as it had the potential to interfere with the demands of the specific tasks in terms of the quantity and quality of information necessary to accomplish goals. *ISp and SS are not the most effective ways to mitigate the influence of ocular dysfunction on computer interaction. *The interactions of SS, ISp, and Sz with ocular factors did not affect the highlight portion of the task, also reflected in their minimal influence on errors. *SS and ISp were more commonly observed, in the models, to effect the interaction consistently across all participants, particularly for the non-clinical models. Changes to these factors in interface design therefore are not deemed critical in the accommodation of needs for individuals with visual impairments.

Table 4.28. continued.

<table>
<tr><td colspan="5">Hypothesis 2: The potentially positive influence of auditory feedback on the drag and drop task is effectively offset by the complexity of the task (multiple icons and multiple targets as compared to the single file – single folder task used in previous studies).</td></tr>
<tr><th rowspan="2">Contributing Results</th><th>Efficiency</th><th>Accuracy</th><th>Information Processing</th><th>Subjective Response</th></tr>
<tr><td>✓</td><td>✓</td><td>✓</td><td>✓</td></tr>
<tr><th>Supporting Evidence</th><td colspan="4">*Positive influence of AF was not observed across all participants. Only TTHT decreased in the presence of AF for the non-clinical model alone, an effect which dissipated in the model when the VSAT score increased.

*The clinical VST model showed signs that VST imposed performance decrements by slower visual search terminations in the conditions that included AF, but contributed to faster times for the intervals with AMD, particularly at the more pronounced levels of severity.

*MT, TTHT, and TTHT were faster in the presence of AF, but only for individuals with AMD, and more so for participants with the most severely diagnosed eyes.

*AF was only observed, in the non-clinical models to interact with VSAT, with increased TTHT into the model when the VSAT*AF term increased in value.

*All participants demonstrated significantly greater mean saccade duration in those trials incorporating AF.

*Participants' overall opinions to AF were mixed.</td></tr>
</table>

Table 4.28. continued.

Conclusions	*Several patterns emerged in the results concerning AF that confirm, but also enriched previous findings. *Jacko and Colleagues (2004, 2005) demonstrated, for a simple drag and drop, that AF was the most helpful modality across participants, and furthermore did not cause any degradation of performance in any participants (including visually healthy). *To the contrary, Vitense (2003) observed detrimental affects of auditory feedback in a complex drag and drop task; performance times were slower in the presence of the auditory cue. The participants in this study were exclusively visually healthy.
	*The current study on the desktop favors the results of both, which suggest that participants must have a visual impairment to realize any benefit from the inclusion of this supplemental non-visual feedback, because AF can pose carryover effects onto other task performance requiring visual attention. *The feedback showed a propensity to impose carry over effects into other non-related components of the interaction, unrelated to the release of the icon, which suggests that AF should not be integrated as a 'quick and dirty' solution to improved performance. Instead, contextual, task, and participant-related factors (including preference for sound) need to be accounted for in the judicious incorporation of AF. *The feedback did improve interaction on the intended component of the task (the drop, as shown with TTHT and MT). However, it only affected the efficiency, not the accuracy in which the movement of the icon was carried out. *The results attributed to AF were largely inconclusive. With a lack of clear, main affects attributable to AF across all participants. *The inclusion AF on this task with multiple files and folders did not equate to the importance of changing icon size for this set of participants.

Table 4.28. continued.

H2a. The presence of auditory feedback will not affect the visual search component of this task				
Contributing Results	**Efficiency**	**Accuracy**	**Information Processing**	**Subjective Response**
	✓	✓	✓	
Supporting Evidence	*The AF*VFQ interaction term emerged as an influential factor in increasing VST in the clinical model *Participants' mean saccade duration was significantly longer in the trials including AF, and this was especially salient for the participants with no ocular disease present.			
Conclusions	*The use of supplemental non-visual cues may in fact be distracting to those individuals with lower or no visual dysfunction, to the point where it interferes with non-related task components. Carry over effects of a sound into other task interactions should be carefully considered. *The increased saccade duration also suggests that AF imposes negative impacts on the visual search component of the task, across all participants. Combined with the result the outcome increased VST for those at higher levels of visual function suggests that the presence of AF redirects the visual scan patterns of the visually apt population in a negative manner.			

Table 4.28. continued.

H2b. As the number of icons and potential drop targets increase, the presence of auditory feedback can have detrimental effects on the interaction.

Contributing Results	**Efficiency**	**Accuracy**	**Information Processing**	**Subjective Response**
	✓	·	✓	·
Supporting Evidence	*Both SS and AF imposed longer visual search times in the clinical model of VST. *Mean saccade duration was significantly longer for participants in the trials including AF.			
Conclusions	*The negative influence of AF in visual search can compound the increased demands on the visual processing system imposed by a larger set of distracter icons.			

Table 4.28. continued.

Hypothesis 3: Measures of ocular health and visual function are predictors of performance in the required task.				
Contributing Results	**Efficiency**	**Accuracy**	**Information Processing**	**Subjective Response**
	✓	✓		✓
Supporting Evidence	*Main effects from the ocular health measures were observed, but were not consistently influential on all aspects of performance, as they were in the handheld task. NVA was the most prevalent influence of performance on desktop interaction efficiency and accuracy. *Increases in AMD Severity (AMD Score) were influential on decreased MT and increased the likelihood for IA errors. *Improved values to CS were observed to impose an increase in the predicted probability for FTHT to exceed the 85th %tile. However, the relative magnitude of this influence was minor compared to the influence of deficits in NVA on increased FTHT.			
	*Diminished NVA (values approaching 1) was observed to monotonically influence performance decrements, mainly in the efficiency measures (TT, VST, MT, TTHT, FTHT), but also imposed an increased likelihood for AD. *Only one main effect was observed for the non-clinically-acquired factors. Increases to VFQ overall score, or improved perceptions of visual function on daily activities, was observed in the model of VST to generate a faster rate of visual search termination *No main effects were observed for VSAT in these models for desktop interaction. *The affinity for the larger icons was shown to be a function of VFQ score. Those with lower VFQ overall scores were more apt to state a preference for the larger icons.			
Conclusions	*Task efficiency on the desktop computer, for this complex drag and drop task, was largely directed by NVA, the ability to focus on the fine details of the screen. *Task accuracy, was largely influenced by interface factors, extraneous personal factors, or interactions between interface and visual factors. However, the influence of visual factors overall was shrouded by the other factors.			

Table 4.28. continued.

<table>
<tr><td colspan="5">H3a. Certain components of the interaction (e.g., the visual search, drag and drop) are more susceptible to the negative impacts of limited ocular functioning.

H3b. The components of the interaction are influenced by each measure of ocular health and visual function to a different degree.</td></tr>
<tr><td rowspan="2">Contributing Results</td><td>Efficiency</td><td>Accuracy</td><td>Information Processing</td><td>Subjective Response</td></tr>
<tr><td>✓</td><td>✓</td><td></td><td></td></tr>
<tr><td>Supporting Evidence</td><td colspan="4">*NVA monotonically affected performance decrements to the efficiency components of the task.

*Neither CS nor AMD Score were highly influential (as main effects) in any of the models.

*None of the visual factors were observed to influence MDC, in either the clinical or non-clinical models.

*Neither VFQ nor VSAT (nor interaction incorporating either) was included in the clinical model of OND.

*Visual factors had a relatively minor influence in the non-clinical models of AD, IA, OND.</td></tr>
<tr><td>Conclusions</td><td colspan="4">*Personal and Interface factors were the major force behind accuracy, as observed in the desktop, in lieu of the ocular factors (with exception of NVA in OND errors). This suggests that while the interactions are slower in the presence of limitations to the visual sensory channel, there is likely a speed accuracy tradeoff for this set of individuals.

*The accuracy of the takes, for the most part, is influenced more by personal and task-related factors, suggesting the extraneous factors that influence mouse manipulation weigh more heavily on the precision of the icon manipulations than vision. In addition, while the effects of NVA appear to be diminished in the presence of changes to Sz and AF, there are few interactions observed between the visual metrics and interface components for the accuracy assessment.</td></tr>
</table>

Table 4.28. continued.

H3c. Different measures of ocular ability can delineate which components of the task will be executed in a less efficient manner, following the speed-accuracy tradeoff common to most HCI tasks. **H3d.** The predictive power of attributes of visual function used to classify the outcome of interactions will differ in terms of the amount of influence on the task.				
Contributing Results	**Efficiency**	**Accuracy**	**Information Processing**	**Subjective Response**
	✓	✓	✓	
Supporting Evidence	*NVA monotonically imposed performance decrements on all efficiency measures; slower TT, VST, MT, TTHT and FTHT as NVA worsened in the models. *The only accuracy measure to demonstrate effects of poor NVA was an increased likelihood for AD. *The presence of AMD and its increase in severity imposed an increased likelihood for IA errors, but actually decreased the likelihood for MT. *The interaction between higher AMD Scores and Increased SS led to decreased TX. *CS influenced an increased likelihood for longer FTHT as the CS score improved. *Increased VFQ generated faster VST. *Increases to the interaction term VFQ*Sz influenced slower VST, higher likelihood for AD, TX, but faster TT. This suggests that the participants with lower VFQ scores could improve their visual search and accuracy with the larger icons, but at the cost of overall increases in TT. *The VSAT*Sz interaction demonstrated that increased values in this term led to slower TT and slower MT. While the magnitude of this influence was minimal, it suggests an important trend underlying the efficiency models. For individuals with higher levels of visual attention, increase in icon size may not be as critical in the improvement of overall efficiency (TT) and in fact may result in interference with the completion of the movement portion of the task. The gains offered by increases in icon size may not be realized for individuals who have higher levels of visual attention. Furthermore, the tradeoff of decreased information presented at the larger icon size may not warrant the use of the larger icon size for certain populations.			

Table 4.28. continued.

Conclusions	*The speed accuracy tradeoffs were not strongly supported in the desktop analyses, but there are underlying trends that support its presence. The fact that the visual factors were not highly influential across all of the accuracy measures, but were across all efficiency measures, suggests an underlying speed accuracy tradeoff for these individuals. *The high accuracy demonstrated by the participants in the desktop task raises question of what would happen in the context of time constraints to the task. It is hypothesized that the participants accuracy levels would falter in the presence of the time constraints, and that those individuals who took the longest time in the present study; those with visual dysfunction and ocular disease, would experience the largest decrements to performance if time limits were imposed.

Table 4.28. continued.

H3e. The influence of ocular functioning on task interaction greatly overrides other normal, age-related declines in mental and physical health.				
Contributing Results	**Efficiency**	**Accuracy**	**Information Processing**	**Subjective Response**
	✓	✓	✓	
Supporting Evidence	*The SF-12 MCS continually emerged as a predictor of performance, both accuracy and efficiency in both clinical and non-clinical models. Increases in perceived mental health were found to improve performance efficiency and accuracy. *SF-12 MCS was one of the most influential predictors included in the models for TT, VST and MT for both clinical and non-clinical approaches. *SF-12 PCS was shown to slow TT in both clinical and non-clinical models, increase VST in the clinical model, and MT in the non-clinical models. *Dexterity demonstrated effects in terms of a speed accuracy tradeoff between searching for and moving; improved dexterity was shown speed up TT and VST; decrease the likelihood for AD and IA, but slow down the MT, TTHT, and FTHT components of the task. It is important to recognize that the slower MT, TTHT, and FTHT did not impose performance decrements on TT. *Age demonstrated a speed accuracy trade-off- MT was longer, but the participants were faster in their search for the icons and were less likely to commit OND and AD as age increased in the model. *Only personal and interface factors were included in the clinical models for OND and MDC, and the non-clinical models for MDC.			
Conclusions	*Clearly, the effects of personal factors are highly influential in the desktop interaction, yet prove challenging to fully account for in determining the efficacy of different adaptations to the interface. *Future studies need to consider how to best capture, and account for extraneous personal factors such as age and aspects of health. *The interactions on the desktop involved a greater degree of coordination than the handheld - especially visual motor coordination, and the ability to project the physical movement of the mouse into movement of the pointer on the display. These components of desktop interaction are highly reliant on factors relevant to physical health, mental ability, and endurance. *Future studies should examine the interactions between these common co-morbidities along with visual factors			

Table 4.28. continued.

<table>
<tr><td colspan="5">H3f. Non-clinically-acquired measures of visual function, such as the users' perception of the impact of visual dysfunction on their activities of daily living (VFQ), and functional visual attention are more powerful predictors of the outcomes of the interaction than clinical factors.</td></tr>
<tr><th rowspan="2">Contributing Results</th><th>Efficiency</th><th>Accuracy</th><th>Information Processing</th><th>Subjective Response</th></tr>
<tr><td>✓</td><td>✓</td><td>✓</td><td>✓</td></tr>
<tr><th>Supporting Evidence</th><td colspan="4">*The outcomes from the VFQ and VSAT models were chiefly dominated by the personal-related, task-related predictor variables.

*There was only one instance where the non-clinical predictors imposed a main effect on the model - increased VFQ generated faster VST.

*The R^2-adjusted values for the non-clinical linear regression models were lower than those generated for the clinical models, which suggests the models with VSAT and VFQ are deficient in fully accounting for the constructs relevant to interaction with the computer – more so than NVA.

*The monotonic relationships between NVA and the efficiency measures suggests that core components of visual function drove the interaction on the desktop, which not captured by other summaries of visual function and diagnosis.

*VFQ was included in the models based on its interaction with Sz, showing a propensity for those with worse VFQ scores to benefit more from increased icon sized in terms of VST, AD, TX and TT. In addition, those individuals who rated the largest icons as most helpful had significantly lower VFQ sores than the group who preferred the standard icon size.

*VSAT did demonstrate how AF was more influential for individuals with lower VSAT percentiles.</td></tr>
<tr><th>Conclusions</th><td colspan="4">*The diagnosticity of the clinical visual factors, for the desktop interaction, was not replicated in the non-clinical measures.

*While they were not as effective in their prediction of performance as the clinically-acquired measures, they were able to capture the efficacy interface modifications, such as increased Sz, in addition to the preference of the participants for Sz. There is more potential in combining VFQ with the clinically acquired models.

*The relationship of VSAT percentile for the desktop with performance provided very little insight between this non-clinically acquired assessment of visual and cognitive functioning and computer interaction. However, the ease with which the VSAT is collected, and the limited resources it requires, compared to the clinically-acquired factors motivates further consideration of this tool

*Future studies should make more direct comparisons between the psychometric properties of the non-clinical and clinical ocular factors.</td></tr>
</table>

4.8 References

American Foundation for the Blind. (2004). "Statement on technology and aging: statement for the record of the American Foundation for the Blind prepared for the United States Senate Special Committee on Aging regarding assistive technologies for independent living: Opportunities and challenges, April 27, 2004 hearing." Retrieved July 8, 2004, from http://www.afb.org.

Backs, R. W. and L. C. Walrath (1992). "Eye movement and pupillary response indices of mental workload during visual search of symbolic displays." Applied Ergonomics **23**: 243-254.

Darling, W. G., J. D. Cooke and S. H. Brown (1989). "Control of simple arm movements in elderly humans." Neurobiology of Aging **10**: 149-157.

Everett, S. P. and M. D. Byrne (2004). "Unintended effects: Varying icon spacing changes users' visual search strategy." CHI Letters **6**(1): 695-702.

Fahrmeir, L. and G. Tutz (1994). Multivariate Statistical Modelling Based on Generalized Linear Models. New York, Springer-Verlag.

Field, A. (2000). Discovering Statistics Using SPSS for Windows. London, Sage Publications.

Hart, S. G. and L. E. Staveland (1988). Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. Human Mental Workload. N. Meshkati. Amsterdam, The Netherlands, Elsevier Science Publishers B. V. (North-Holland)**:** 139-183.

Jacko, J. A., K. P. Moloney, T. Kongnakorn, L. Barnard, P. J. Edwards, V. K. Leonard, F. Sainfort and I. U. Scott (2005). "Multimodal feedback as a solution to ocular disease-based user performance decrements in the absence of functional visual loss." International Journal of Human-Computer Interaction **18**(1): 183-218.

Miller, S. M. (1987). "Monitoring and blunting: validation of a questionnaire to assess styles of information seeking under threat." Journal of Personally and Social Psychology **52**(2): 345-353.

Neter, J., M. H. Kutner, C. J. Nachtsheim and W. Wasserman (1996). Applied Linear Statistical Models. Boston, McGraw-Hill.

Pallant, J. (2003). SPSS Survival Manual. Philadelphia, Open University Press.

Rogers, W. A. (1997). Individual differences, aging, and human factors: An overview. Handbook of Human Factors and the Older Adult. A. D. Fisk and W. A. Rogers. London, Academic Press**:** 151-170.

Smith, M. W., J. Sharit and S. J. Czaja (1999). "Aging, motor control and the performance of computer related mouse tasks." Human Factors **41**(3): 389-396.

CHAPTER 5

RESULTS SUMMARY

5.1. Introduction

This chapter serves to synthesize the results presented in Chapter 3 and Chapter 4, highlighting the most compelling contributions posed by this thesis to the existing body of work. Comparisons are also drawn between the outcomes of the desktop and handheld experimental tasks, making conclusions concerning Hypothesis 4. The relative contributions of the thesis to the existing body of work in HCI and visual impairment, and general HCI are also presented. The chapter concludes with a collection of recommendations for future analyses and empirical exploration of the subject area, targeting the further development of the framework of interaction thresholds.

5.2. Handheld & Desktop Contrasts: Hypothesis 4

Hypothesis 4, first introduced in Chapter 2, considered how the two platforms, handheld and desktop, would differ in terms of the affected interactions and resource requirements. However, because the experiment was blocked on the platform condition, a different set of participants worked with the desktop versus the handheld. This design was used mainly to limit any learning effects between the two platforms, but also because of limited experimental resources, and fatigue of the participants. Because of the high levels of variability inherent to this older population with visual impairments, the conclusions drawn from comparisons between the desktop and handheld PCs are necessarily limited. To perform a more complete comparison would require the same set of participants to work with both settings, and/or a substantially large sample of participants to evoke the principles of the central limit theorem. Another factor that constrains the comparisons between desktop and handheld is the difference in the number of repetitions completed by each participant under experimental conditions. The desktop repetitions were fewer in number, due to the inclusion of the additional independent variable, icon size (Sz). Hypothesis 4 and its sub-hypotheses are listed below, with results and discussion relevant to each one.

Hypothesis 4:

> *The interactions of those users in the handheld PC group will be less efficient than those users in the desktop PC group*.

The overall efficiency of the groups on the tasks was comparable. The total time to complete the task for the handheld averaged 19.71 minutes and for the desktop was 15.31 minutes. Shown in Figure 5.1, the mean time spent in the various components of the task was similar. The desktop task took longer in their time per trial and in the movement time of the card icon from its original location to the drop pile (and not unexpectedly). The additional physical space, which defined the visual search and movement of the mouse on the desktop, was considerably greater in size.

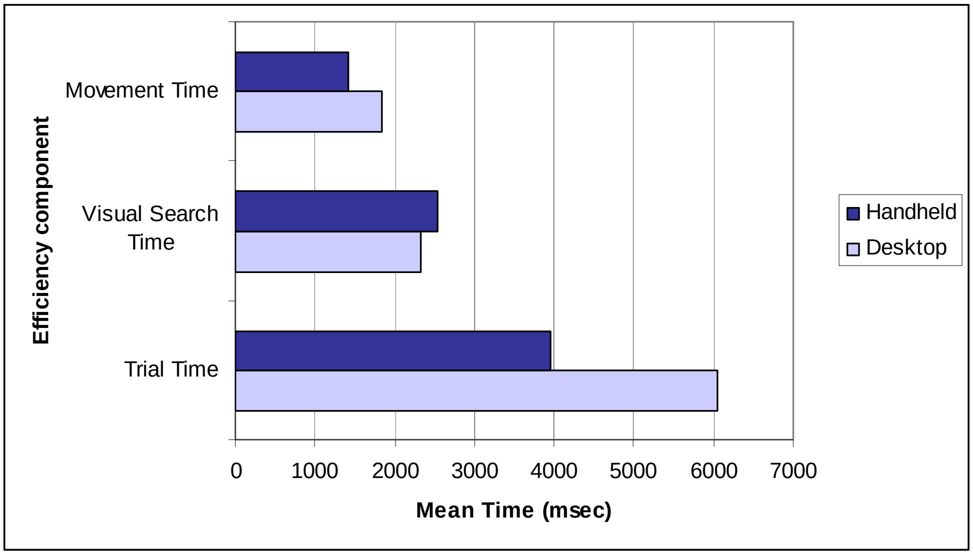


Figure 5.1. Comparison of efficiency measures between handheld and desktop platform experiments.

The visual search time was slightly longer on the handheld, not surprisingly so, with all trials subject to the smallest icon size in the handheld condition, 7mmx7mm.

The time spent in visual search relative to icon movement for each platform is graphically depicted in Figure 5.2. For the participants who interacted with the desktop, the ratio of time spent in visual search versus icon movement was closer to one, than it was for the handheld group. In the desktop, this is rationalized by the larger area over which the icon had to be moved in terms of the physical space on the display relative to the required arm and hand movements. It required the participants to integrate a combination of both gross and highly precise arm and hand movements with the mouse as they visually tracked its icon and cursor movement across the display. This suggests potential implications attributable to both input device and display size. The use of the stylus for the handheld input removed the additional layer of abstraction that mouse requires. For mouse input, users must effectively integrate their movement in the physical world with the movement they perceive in the cursor on the display. Alternatively for stylus-based movement, the distance the stylus moves reflects the absolute distance the icon will move on the display.

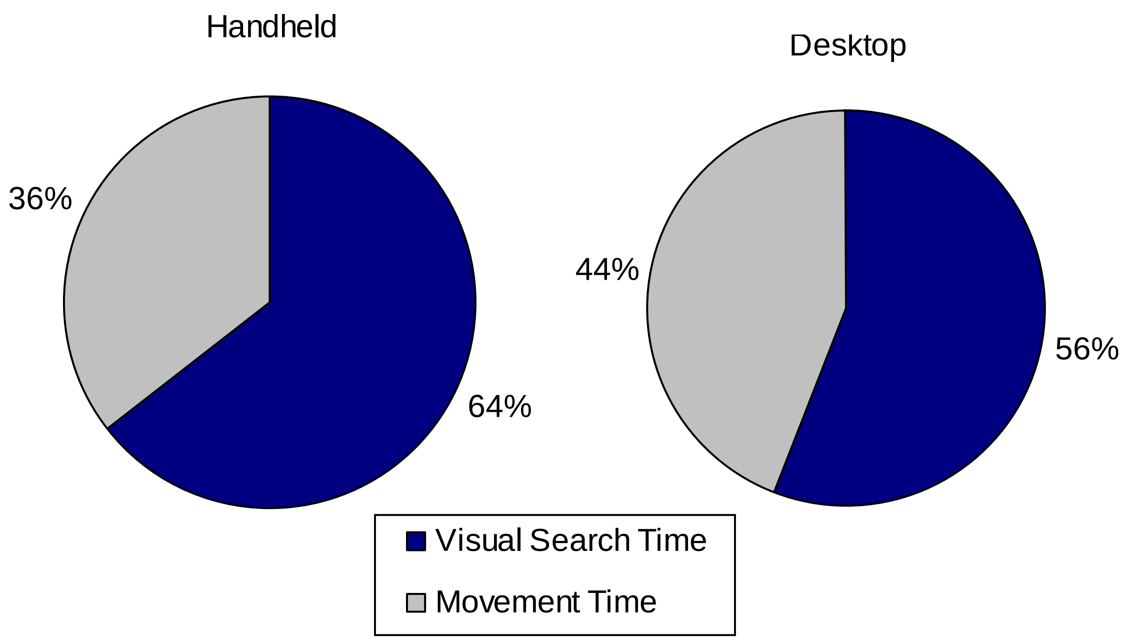


Figure 5.2. Percentage of trial time spent in visual search versus movement time according to platform, desktop or handheld PC.

Hypothesis 4a & 4d:

> *The motor skill required by the input device will cause users to slow their performance at different points in the interaction, demonstrating a speed–accuracy tradeoff not readily observable in the desktop PC condition; Normal age-related declines to mental and physical health will be amplified by the interaction style required by the handheld PC.*

While visual factors were included as the most dominant predictors across all phases of interaction for the handheld, this was not the case in the desktop models. In the desktop, several of the models were driven by non-visual factors. In particular, these factors were indicative of the constructs related to normal age-related declines in mental and physical health. In addition over no drop errors and movement direction change errors were observed in the desktop to not be influenced significantly by visual factors, but instead interface features such as target location, set size and icon size.

For the desktop, SF-12 PCS and SF-12 MCS, and were highly influential in the both clinical and non-clinical models on all of the time measures (trial, visual search, movement and highlight) and the probability for accidental drops of the card icon analyses. The SF-12 scores were monotonically related to both accuracy and efficiency models – as the scores on mental and/or physical health decreased (worse health states) the likelihood for errors increased, and the rate of completion of the assorted task phases was slower.

Again, this is reflective of the requirements the desktop computer places on the individuals, and how it is a more involved interaction on many levels. The demands placed on the various resources, physically, and cognitively, are greater. The additional abstraction and physical movement required by using the mouse demands a greater number of resources to operate in coordination, than does the use of the stylus.

In terms of manual dexterity and motor control, the critical role manual dexterity plays in computer interaction was demonstrated in both desktop and handheld experiments. A speed accuracy trade-off was observed for the desktop relative to dexterity. Improvements to manual dexterity induced faster trial times, visual search times and resulted in a decreased probability for icon acquisition errors and accidental drops (for both clinical and non-clinical). The trade-off was observed in the 'drop' phase of the task, where the TTHT, FTHT and MT were all slower as dexterity improved. In other words, those with worse dexterity were faster with their movements of the mouse, but experienced greater difficult in the selection of the card icon and were prone to drop the card before reaching the drop pile. For the handheld, a similar trend was observed as improvements in dexterity increased movement time, but decreased trial time, drag distance, the likelihood for accidental drops, task axis crossings, and movement directions changes.

A surprising and noteworthy result emerged based on the inclusion of dexterity in the models on performance. The inclusion of dexterity was consistently included in all four models of visual search time. As dexterity improved, visual search time improved in clinical and non-clinical models for both the desktop and handheld. This result suggests a link between mouse movement and visual search. Reviewing video clips of the participant interaction showed several participants who used the mouse pointer or the stylus to guide/keep their place during their visual scan of the display. The pointer or stylus served as a place marker on the screen. A deficiency in manual dexterity could effectively inhibit this coordinated task between the visual and motor systems.

The significance of this result affects the future of computational modeling in HCI. There is some debate in the computational modeling community concerning the measure of visual search (Byrne 1993; Everett and Byrne 2004). Some researchers do not allow the users to move the mouse until they locate the cursor and treat them as discrete components of the interactions. In many experiments for computational models and HCI, the participants are prohibited from moving the mouse until they have visually located a target. In addition, the resultant models reflect these distinct components.

This trend suggests the interconnection between dexterity and visual search. While it does not suggest that all users couple their mouse movement with visual search, it strongly alludes to the complexity of interactions between the psychomotor and visual sensory systems for those who do employ this technique during visual search. The exclusion of the effects of the coordination of motor and sensory function in visual search may decisively limit the applicability of the results produced by computational models.

This phenomenon should be further explored and isolated in future empirical work. Specifically, the tendency to use a pointer device to guide visual search merits exploration. It could be that this phenomenon is exaggerated in the aging population, and those with visual impairments. The ability to complete efficient searches, filtering out extraneous noise and distracters has been observed to decrease with age (Owsley and Sloan 1990; Schieber 1994; Orr 1998). It is possible that the use of the mouse pointer or stylus serves to direct attention during visual search. Additionally, this trend should be investigated for different types of pointing devices and cursors. The results of such a study could inform the design of interface input devices to best accommodate individuals who experience competing demands for attention resources. This effect is likely amplified by the age of the participants in this study. The faculty to share resources, attention and psychomotor, has been shown to decline with age – the different channels are more prone to be overloaded in the presence of distracters. A more direct, less consuming interaction such as the use of the stylus, or pen based interfaces, is more amenable to the general aging population.

Hypothesis 4b:

> *Users with visual impairments will experience more performance decrements with respect to interactions on a handheld device than the visually healthy cohorts.*

Visual impairment played a key role in the classification of interactions on both the handheld and desktop. While the desktop was observed to be more affected by personal and interface features than the handheld, visual factors were observed to be the considerable driving force on several of the models for both. A single visual factor served as the most salient influential predictor in both the desktop and the handheld, but the factor was distinct to each platform.

A universal monotonic trend in diminished performance was observed for performance in both handheld and desktop conditions, attributable to a clinically-acquired ocular factor. For the handheld, AMD Score, the severity level of AMD, was shown to be a significant predictor of both accuracy and efficiency. For the desktop condition, decrements to near visual acuity were observed to be the most common predictor to substantially compromise efficiency and accuracy. Tables 5.1 and 5.2, which first appeared in the outcome summaries for Chapters 3 and 4, capture the breadth of impact that AMD Score and near visual acuity had over the respective platforms.

Table 5.1. Desktop PC impact summary for ocular factors, both clinical and non-clinical.

Clinical	**Predictor: AMD Score** Increased disease severity level (worse)	
	Increasing	IA
	Decreasing	MT
	Predictor: CS Score Higher, improved contrast sensitivity	
	Increasing	FTHT
	Predictor: NVA Score Decrease in visual acuity, worse vision	
	Increasing	TT
		VST
		MT
		TTHT
		FTHT
		AD
Non-Clinical	**Predictor: VFQ Overall** Improved perception of visual function	
	Decreasing	VST
	Predictor: VSAT Percentile Higher percentile for visual attention (improved)	
	None	

Table 5.2. Handheld PC impact summary for ocular factors.

Clinical	**Predictor: AMD Score** Increased disease severity level (worse)		
	Increasing	TT	
		VST	
		MT	
		DD	
		TTHT	
		OND	
		AD	
		TX	
		MDC	
	Predictor: CS Score Higher, improved contrast sensitivity		
	Increasing	DD	
		AD	
		MDC	
	Decreasing	TT	
		VST	
		MT	
		Pupillary Change	
		Fixation Duration	
	Predictor: NVA Score Decrease in visual acuity, worse vision		
	Increasing	TT	
		MT	
		TTHT	
		OND	
		TX	
	Decreasing	DD	
		AD	
		MDC	

Table 5.2. continued.

Non-Clinical	**Predictor: VFQ Overall** Improved perception of visual function	
	Decreasing	DD
		MDC
		AD
		TX
	Predictor: VSAT Percentile Higher percentile for visual attention (improved)	
	Decreasing	MT
		VST
		TT

The difference in the strength of AMD versus NVA is justified between the two different platforms. AMD severity measures the number of drusen on the macula and those with the most severe levels of AMD in an exudative (wet) state. The patients whose eyes show the presence of drusen and/or the presence of wet AMD are more prone to face the types of functional declines most common to AMD. These dysfunctions commonly include interruptions, distortions, and loss of fine detail vision in the center of their visual field.

The difference in the influence of AMD severity level on the handheld versus the desktop is attributable to the difference in physical display size. That is, the disturbances in the visual field will occlude a larger percentage of the handheld display than they would block on the desktop. The effect of visual field disruptions is more problematic for the perception of the handheld display.

This difference can be proven in considerations of visual field, viewing distance, and display size. Assuming a viewing distance of 40 cm on average for both the handheld and desktop and using the estimation that human binocular vision is 180 degrees (optimistic for an aging population) the visible field at once is an area roughly 246 in^2. This is more than 10 times the area of the handheld display but accounts for just 68% of the desktop monitor. This indicates that *none* of the participants could perceive the entire desktop display at one time (while the specific amount changed). However, for those participants on the handheld, the visually healthy controls could easily account for the entire handheld display in their visual field.

This being the case, for those individuals with AMD, as severity increased, they were increasingly prone to experience interruptions to the visual field. In the presence of consistent interruptions and distortions to the visual field, a greater number of card icons and visible features would be distorted or even absent form view on the handheld, than with the use of the desktop display. Figure 5.3 illustrates the different impact of the same interruption in visual field on the handheld versus the desktop. A visual field interruption consistent in size and shape was applied to both images. As illustrated, a larger percentage of the display is interrupted for the handheld. The affects of the increasing AMD severity are amplified on the handheld, observed in the analyses of the efficiency and accuracy outcome measures. This demonstrates the effects of a small display for a disease such as AMD.

Despite the measured affects of AMD severity level on the handheld, it is remarkable that all of the participants but one completed all of the trials. These individuals were able to work with the technology, (overall successfully) in spite of their disease. Furthermore, the observed gains in performance measured in the interaction terms of AMD and AF provide an easily implemented, low cost solution to mitigate the effects of the disease. Research and commercial efforts both need to address the needs of aging adults in the emergence of new technologies, as they have great potential, are have been demonstrated to be usable. Furthermore, as will be shown in Chapter 6, this segment of the population is growing, will be actively seeking ways to maintain active lifestyles, and will have significant purchasing power in future markets, especially as the as the Baby Boomer population advances in age.

Figure 5.3. Illustration of the relative impact of AMD on handheld versus desktop, for moderate stages of AMD visual field disruption consistent between the desktop and handheld image (displays not to scale) (Image simulated from MD Simulator 3.1, www.opticaldiagnostics.com).

The fact that NVA dominated the models of performance on the desktop, in terms of both efficiency and accuracy, is not surprising. Near visual acuity has been observed to be a valuable predictor of performance for patients with AMD in several other studies of HCI and visual impairment (Jacko 1999, 2000). In addition, several studies have stratified groups of AMD patients based on visual acuity and observed performance differentials according to these groups, for a less complex, but similar task on a desktop. AMD, at more pronounced stages affects the fine detail resolution of central vision, making near vision task, such as reading, sewing, or using a computer especially challenging (Macular Degeneration Partnership; Partnership 2005). Its strong influence is therefore not surprising, as the task was 1) Visually intensive (as shown in Figures 5.1 and 5.2) and 2) required a significant level of visual resolution to detect differences between both the alphanumeric characters and suit shapes used with the playing cards. Also, it should be noted that near visual acuity, for the handheld, was the most influential clinical ocular factor compared to all other predictors.

The percentage of the display viewable at one time also lends explanation to the interesting effects of set size observed in the handheld PC interactions, but not on the desktop. For the handheld, increases in set size were observed to slow the rate of VST and TT across all participants, as a main effect. However, there was an interaction effect, based on the term AMD*SS. In the presence of the disease, the negative effects of increased set size were essentially nullified. Patients with higher AMD Scores were diagnosed with eye(s) at more advanced stages of the disease. This meant a larger number of drusen on the macula, and even observed leakage (wet AMD) at the higher AMD Scores.

Figure 5.4. Illustration of set size effect for participants diagnosed with AMD, (Image simulated from MD Simulator 3.1, www.opticaldiagnostics.com).

For participants experiencing disruptions to the visual field, it is likely for a change in the number of icons presented on the screen at one time to not be perceived as different. That is, they may not even notice the change in the number of distracters at the onset of the display, and, unlike the participants with normal, clear vision, may not adjust their visual search strategies with the presentation of the grid of card icons. The aberrations experienced within the visual field and the strategies used in the initial visual scan of the display were likely consistent between trials, independent of the number of icons appearing on the display at stimulus onset. Figure 5.4 demonstrates the phenomenon between change in set size and ocular diagnosis. This is more likely an observable occurrence on the handheld because of the display size. Because the entire desktop display is not visible in one visual field, the extent to which the initial scan patterns differ according to disease severity is not as extreme.

Hypothesis 4c:

> *The most detrimental component of the handheld PC to interactions for individuals with visual impairments is the size of the icons.*

While the handheld employed only the smallest icon size, 7mmx7mm, *none* of the participants on the handheld task requested to skip trials due to this fact, and *none* of them complained about the icon size (even in the exit interviews). This is curious, when juxtaposed with the number and characterization of the skipped trials by participants in the desktop conditions. In the desktop condition, skipped trials were largely attributable to the smallest icon size, and participants voiced a preference for the largest icon size in the exit interviews. The participants, as shown in Chapter 2, were statistically equivalent in terms of their ocular and personal factors between desktop and handheld.

In addition, there was a large measurable effect of icon size in the models for the desktop. In terms of the display, for those conditions using the smallest icons, the only difference was the amount of white space (or black space in this experiment) on the display and the mouse cursor. The implications of display size should be explored to further clarify what drives the differential in the ability and the perception of capability to complete the task. In our experiment, participants in the desktop setting were more easily overwhelmed in the presence of the smallest icon size, to the point where tasks were skipped.

The effect of size on the desktop is also interesting, in terms of previous findings. Previous work with icon size pointed out a distinct point of diminishing returns with increases in icon size, which was not readily observed in this dissertation experiment (Emery, Jacko, Kongnakorn et al. 2001). In addition, several others warn of the possible occlusion of icons larger in size for those individuals who experience interruptions in their visual field (e.g., Kline and Glinert 1995). That said, a possible explanation for the 'bigger is better' effect observed for the desktop interaction is the features of the icons included in the task. The card icons share many similar features, and are distinguished by slight changes in shape (suit), color, or number. This is different than application or functional icons, such Microsoft Office icons used by Jacko and colleagues (e.g., a printer, a document, a garbage can, a folder, etc.) (e.g., Jacko, Barreto, Chu et al. 2000; Emery, Jacko et al. 2001; Vitense, Jacko and Emery 2003; Jacko, Moloney, Kongnakorn et al. 2005). In that experiment, the basic shape of the icons could deviate substantially. The card icons, on the other hand, were much more similar in nature to each other and distinguished by fine alphanumeric qualities and distinctions between shape and color. This is also a likely contributing factor to the dominant role near visual acuity took in the task – the ability to discern fine details, such as the differences between a club and a spade, or a number ‘3’ versus a number ‘8’ would necessitate a sufficient level of visual acuity, at a distance near to the computer display. In fact, on the handheld, NVA was often the second most influential factor of performance, which can also be attributed to the need to extract distinguishing features from the card icons.

5.3. <u>Contributions Summary</u>

The main objective of this thesis was to contribute substantially to the growing body of literature, which informs a framework of interaction thresholds for individuals with visual impairments. The evolution and growth of this framework will eventually serve as tool from which universally accessible design guidelines can be derived. Table 5.3, first introduced in Chapter 1, and has been updated to demonstrate the contributions posed by the outcomes of this work. As shown by the darker boxes highlighted in the bottom row, the contributions of this dissertation present incremental increases in terms of existing knowledge on HCI and visual impairment. In addition, the dissertation outcomes provide substantial advancements to the knowledge base in terms of its increased level of both contextual and task validity beyond previous work. This section highlights the most influential outcomes of this work that factor into the expansion of the existing knowledgebase.

It is important to consider how the outcomes of this work emerged to corroborate previous findings, or if they established new trends divergent from the previous conclusions. The results also extend the knowledge base for existing HCI research. These outcomes relate to pervious studies of icon use, display density and auditory feedback with the drag and drop. Table 5.4 summarizes for the domain of HCI and visual impairment the existing studies, conclusions, and whether or not the conclusions were supported by the current dissertation. Table 5.5 discloses how the thesis results compared against existing research. For both Tables 5.4 and 5.5, the column marked *Yes/No*, discloses the extent to which the current results substantiate the existing conclusions. The *Justification* column provides evidence from this dissertation corroborating the answer in preceding column.

Table 5.3. Summary of the contributions of this thesis relative to the existing knowledge base in HCI and visual impairment.

	Visual Impairment			**Age**		**Task Phases**					**Com-plexity**		**Independent Variables**						
Citation	Visually Healthy	AMD	DR	General Population	Older Adults	Visual Search	Object Selection	Cursor Movement	Drag & Drop	Menu Selection	Simple	Complex	Icon Size	Set Size	Background Color	Clinicall Visual Profiles	Non-ClinicalVisual Profiles	Multimodal Feedback	Text Size
Jacko 2000; Jacko Barretto, Scott, Rosa & Pappas, 2000; Jacko et al., 1999; Jacko & Sears 1998		■	■	■	■	■	■				■	■	■	■	■				
Jacko, Barreto et al., 2002; Jacko et al., 2001; Scott, Feuer & Jacko 2002a Scott Feuer & Jacko 2002b		■			■	■	■				■	■	■	■	■	■			
Jacko et. al, 2000		■			■			■				■	■	■	■	■		■	
Vitense, Jacko and Emery, 2002; Vitense, Jacko and Emery, 2003	■			■					■			■						■	
Jacko et al., 2002c; Jacko et al., 2003; Jacko et al., 2004a Jacko et al., 2004b; Jacko et al, 2005 Leonard et. al, 2005		■			■				■		■					■	■	■	
Emery et al, 2004; Jacko et al. 2004c	■				■				■		■							■	
Edwards et al., 2004; Edwards et al., 2005 Jacko et al, in press			■	■	■	■				■		■	■		■	■	■	■	■
Leonard, 2005*		■			■	■	■	■	■			■	■	■		■	■	■	

Table 5.4. Summary of research on GUI manipulations and visual impairment and corroboration by thesis.

Study and Citation	Study Conclusions	Supported by Current Thesis Results?	
		Yes/No	Justification
Visual search and selection for a range of visual impairment (Jacko and Sears 1998; Jacko, Rosa, Scott, Pappas and Dixon 1999; Jacko 2000; Jacko, Barreto, Marmet et al. 2000)	Contrast sensitivity and color perception were found to be significant when predicting performance time, while visual acuity, contrast sensitivity, visual field, and color perception were all found to have a significant effect on icon selection accuracy.	Yes	▪ Near visual acuity, contrast sensitivity and diagnosis were all included as significant predictors in the models for both handheld and desktop interactions. ▪ Monotonic trends were observed on both time and accuracy for the handheld according to AMD severity, while on the desktop with near visual acuity. ▪ The distinction between visual factors to better predict accuracy versus performance was not observed in these studies, especially on the desktop. ▪ A speed accuracy trade-off was observed for decrements in contrast sensitivity and near visual acuity on the handheld; these individuals took longer, but would be modeled with more accurate movement with the icon.
	Icon size, set size, and background color significantly influence performance.	Yes	▪ The influence of icon size and set size in this experiment was dependent on the phase of interaction; visual search, or icon manipulation and movement. ▪ Increases in set size caused slower search times and trial times, but on the handheld, this trend was not present for the individuals with AMD, especially as severity worsened. ▪ Increased spacing slowed down visual search and movement time and overall trial time in both platforms, but facilitated a more accurate drop, decreasing errors and TTHT.
	The increased time required by individuals with visual impairments to search is not due to delayed engagement, but to time spent in active search.	Yes	▪ Analysis of eye movements did not reveal significant differences based on AMD diagnosis. ▪ Longer mean saccade durations and greater pupillary response were linked to lower contrast sensitivity, indicative of more 'active' visual search at lower levels of contrast sensitivity. ▪ The group with the largest pupillary response also had significantly worse contrast sensitivity – suggesting a higher level of workload in the visual task.

Table 5.4. continued.

Study and Citation	Study Conclusions	Supported by Current Thesis Results?	
		Yes/No	Justification
Visual search and selection for individuals with AMD (Jacko, Scott, Barreto et al. 2001; Jacko, Barreto, Scott et al. 2002; Scott, Feuer and Jacko 2002, 2002)	Visual acuity, contrast sensitivity, weighted average visual acuity, contrast sensitivity and color perception deficits were found to be significantly associated with the performance of search tasks for users with AMD.	Yes	▪ On the desktop, near visual acuity was the most dominant visual factor to influence interaction in the search component, and the movement phases. ▪ Contrast sensitivity and near visual acuity were highly influential in causing participants to delay their interactions in order to attain increased levels of accuracy.
	Weighted average contrast sensitivity was the most sensitive indicator of performance.	Yes And No	▪ For this population, contrast sensitivity was the least sensitive predictor of performance in the desktop interactions. ▪ The handheld demonstrated a significant, interesting effect attributed to contrast sensitivity; a trade off was observed with poor levels of contrast sensitivity triggering longer visual search times and movement times, but achieving higher levels of accuracy. ▪ For the handheld, those participants with the longest saccade duration and largest pupillary response over the course of the task had significantly worse contrast sensitivity.
	Analyzing eye movements confirmed there were differences due to AMD in the visual search and the interface features of icon size, background color, and set size.	Yes And No	▪ None of the observed trends in the eye data were attributable to AMD severity level, and diagnosis. ▪ On the desktop differences in mean saccade duration were significant across all participants in the presence of auditory feedback. In addition, while not significant, this trend was substantially amplified in the control group. ▪ For the handheld, increased fixation duration and pupillary response were linked with decrements to contrast sensitivity. Diminished contrast sensitivity in the visual field is a known affect of AMD. ▪ SF-12 MCS was observed to direct the saccade duration on the handheld, and the saccade to fixation ratio. The better the mental health score, the shorter the duration, and the smaller the value of saccades to fixations.

Table 5.4. continued.

Study and Citation	Study Conclusions	Supported by Current Thesis Results?	
		Yes/No	Justification
Visual search and selection for individuals with AMD (Jacko, Scott et al. 2001; Jacko, Barreto et al. 2002; Scott, Feuer et al. 2002, 2002)	Changing key screen features such as icon size, background color, and set size can improve performance for individuals with AMD.	Yes	▪ The influence of Icon size and set size in this experiment was dependent on the phase of interaction; visual search, or icon manipulation and movement. ▪ Increases in set size affected slower search times and trial times, but on the handheld, this trend was not present for the individuals with AMD, especially as severity worsened. ▪ Increased spacing slowed visual search and movement time and overall trial time in both platforms, but facilitated a more accurate drop, decreasing errors and TTHT.
Cursor Movement for individuals with AMD (Jacko, Barreto, Marmet et al. 2000)	Cursor movement time and velocity were significantly worse for individuals with AMD, and worsened in conjunction with decrements in visual acuity.	Yes	▪ In both desktop and handheld, MT was slower in the presence of visual dysfunction, and especially in the presence of AMD. ▪ While slower with movement of the cursor, the participants on the handheld demonstrated a speed accuracy trade off relative to near visual acuity and contrast sensitivity. Worse visual dysfunction influenced slower times, but accuracy was not compromised.

Table 5.4. continued.

Study and Citation	Study Conclusions	Supported by Current Thesis Results?	
		Yes/ No	Justification
Drag and Drop for individuals with Diabetic Retinopathy (Jacko, Scott, Sainfort et al. 2002; Jacko, Scott, Sainfort et al. 2003; Jacko, Barnard, Kongnakorn et al. 2004; Jacko, Moloney, Kongnakorn et al. in press)	Performance improvements were observed for both visually healthy and AMD users due to the implementation of non-visual/multimodal feedback.	No	▪ In the context of the complex task, potential decrements to performance were observed, for the visually healthy, in the presence of AF. ▪ VST increases in the desktop setting in the presence of AF, and mean saccade duration, were longer in the AF present conditions. ▪ For the handheld, there were few indications that AF was helpful to the visually healthy population.
	The performance gains for the utilization of non-visual, multimodal feedback were greater in magnitude for users with AMD.	Yes	▪ The performance gains attributable to AF by the AMD group were substantial, and provide a simple, inexpensive solution to minimizing the performance differential imposed by visual dysfunction on both the handheld and the desktop. ▪ In most cases on the group with visual dysfunction were observed to improve performance with the inclusion of AF.
	The presence of AMD significantly inhibited user performance, independent of other ocular functions (e.g., acuity, contrast sensitivity, color perception).	Yes	▪ This was observed on the handheld, exclusively. The presence of AMD monotonically influenced performance decrements in terms of both efficiency and accuracy. ▪ The presence of AMD also interacted in surprising ways with increases to the set size. Efficiency was not detracted for the AMD group, like it was for the visually healthy, in the presence of a greater number of distracters on the screen.
Drop Down Menu Selection for individuals with Diabetic Retinopathy (Edwards, Barnard, Emery et al. 2004; Edwards, Barnard, Emery et al. in press)	Multimodal feedback was found less effective than visual enhancements to the selection of items from drop down menus. Ocular factors such as acuity, contrast sensitivity, and visual field were found to impact performance.	Yes	▪ Across all participants on the desktop, the AF was clearly less effective than the change in icon size, both in terms of its influence on efficiency and accuracy, but also eye movements, and the participants' interface preferences. ▪ Changes to the visual display were often the most dominant predictor of the error likelihood models for the desktop. ▪ The efficacy of the feedback in counteracting visual dysfunction was stronger than the visual changes to the interface in both desktop and handheld.

Table 5.5. Summary of research on GUI research and conformation by current dissertation.

Topic and Citation	Study Conclusions	Supported by Current Thesis Results?	
		Yes/No	Justification
Drag and Drop and Multimodal feedback for visually healthy young adults (Vitense, Jacko and Emery 2002; Vitense, Jacko et al. 2003)	Multimodal feedback improves the performance of fully sighted users in a complex search and selection task, but that the inclusion of auditory feedback alone can decrease task efficiency.	Yes	▪ The desktop condition demonstrated a propensity for AF to trigger performance decrements across participants; Saccade duration was longer for all participants when AF was present.
Icon spacing, set size and quality effects on visual search (Byrne 1993; Everett and Byrne 2004)	Larger set sizes diminished search efficiency. The authors concluded that spacing and quality can induce users to employ suboptimal search strategies. Decreased spacing serves to distract, but spacing beyond what is viewable in the visual field at one time, triggers longer times as well	No	▪ There were no negative effects on visual search attributable to decreased spacing between icons, as demonstrated by both Everett and Byrne and Hornoff. ▪ VST was monotonically related, in both the handheld and the desktop, to slower TT and VST. ▪ Increased spacing was observed to facilitate a more accurate drop, in both handheld and desktop interactions. ▪ OND and longer TTHT were less likely when spacing was larger between icons. ▪ The set size effect was not observed in individuals with AMD on the handheld. ▪ Dexterity was linked to visual search, suggesting computational models, which isolate mouse movement apart from visual search to be inaccurate in representation, especially for older adult who use cues to direct and focus their attentional resources.
Visual search and icon selection in two dimensional hierarchies (Hornoff 2001)	Users exhibit a speed accuracy trade off when the space between targets varies. Slower, more accurate interaction strategies were used as the spacing between icons decreased.		

5.4. <u>Conclusions Summary</u>

This section details ten of the most compelling conclusions to emerge from this research. Each conclusion is presented, along with supporting evidence, and the major implications of the outcome in terms of design and future research are highlighted.

Conclusion #1:

The judicious use of auditory feedback can mitigate the affects of visual dysfunction on computer interaction

Auditory feedback can substantially offset the performance decrements imposed by visual dysfunction such as those associated with aging and/or AMD. However, the limitations in the ability to divide attention and filter out extraneous information can induce supplemental non-visual cues to trigger performance decrements for those who are visually healthy. This also demonstrates that solutions for those with visual impairments and ocular disease need to account for the associated mental, physical abilities of the target population, particularly for aging populations.

<u>Supporting Evidence</u>: On the desktop, those participants without AMD were modeled to generate significantly slower visual search times and longer mean saccade durations in the presence of auditory feedback (AF). This suggested the presence of carryover effects from AF into the visual search component of the task, which actually distracted the participants. In studies of visual attention with older

adults, the older population was more apt to exhibit longer reaction times due to the need to divide attention (Tun and Wingfield 1997). Older adults are prone to difficulties ignoring “task irrelevant information,” (Schieber 1994 p. 17). Research also suggests that visual search is slower and less effective for older adults due to a shrinking of the “useful field of view” to which attention can be simultaneously allocated. The size of the useful field of view, for older adults, is especially susceptible to context related factors, such as complexity and cognitive task load (Schieber 1994).

However, the inclusion of AF was shown to consistently interact with the clinical and non-clinical visual factors, narrowing the performance gap for those who experienced visual dysfunction and/or disease. This suggests that the AF was likely extraneous information to the visually health users, in the context of the desktop, and was a more consuming interaction. The demands exceeded their capacity to divide attention and filter out noise, which triggered inefficient visual search patterns.

Participants were mixed in their reaction to the feedback, possibly indicative of the likelihood for efficacy of the sound in the task. In the handheld experiment, the participants who rated the feedback as *helpful* or *very helpful* had a significantly lower VFQ score. Their willingness to consider augmentations is a function of the normal difficulty experienced in daily, visually intensive activities. This may also signify their ability to effectively integrate the supplemental non-visual cue into their interactions.

In addition, the effect of task validity was likely a driving force behind this emergent result. Previous research with the auditory feedback and drag and drop isolated the interaction for a simple drag and drop (single target icon and single target location). That said, it was completely removed from the more rigorous level of visual search required in this task and the presence of distracters. This speaks to the importance of validating universal design improvements in the context of the entire task. In this instance carry over effects were observed from the presence of auditory feedback in the drop/movement phase that negatively affected the visual search phase, particularly for those who possessed higher levels of visual function.

Major Implication: Auditory feedback is a low cost, easily implemented alternative. However, the widespread integration should be carefully considered, especially for older adults. A easily implemented solution to this is the inclusion of 'optional' auditory feedback in direct manipulation interfaces, which users can switch it off if they prefer. Empirical studies, which consider assistive or universal technologies and solutions for access, must consider the interaction in the context of the entire task. This should be a consideration even for those interactions that are seemingly unrelated to the modifications and design changes being considered, because the effects of the interface can have carry over effects into other peripherally related task components.

Conclusion #2:

Mouse/pointer movement is interrelated with visual search mechanisms.

The role of manual dexterity and hand-eye coordination in direct manipulation tasks has been long acknowledged in the HCI community. However, the role of manual dexterity, specifically movement of the input device during visual search, is a somewhat contentious issue. In computational HCI modeling, there is some debate as to how to discriminate between motor movements with the mouse and visual search, and whether the two events are distinct or coupled (Everett and Byrne 2004). In fact, in their empirical work to derive models many researchers do not allow for the movement of the mouse until visual search has terminated. In many experiments for computational models and HCI, the participants are prohibited from moving the mouse until they have visually located a target. In addition, the resultant models reflect these distinct components, not accounting for the movement time and motor skill, until after visual search termination. An unexpected result emerged in this dissertation, as dexterity was observed to implicate visual search time, suggesting a more complex relationship between the two, than what is commonly accounted for in computational modeling.

Supporting Evidence: In both the handheld and desktop results for VST, as dexterity improved, visual search time improved within clinical and non-clinical models. This result suggests that there is a link between mouse movement and visual search. An informal review of the video record also showed the tendency for participants to use the mouse pointer or stylus to guide their search, serving as a place holder. A deficiency in manual dexterity could effectively inhibit this coordinated task between the visual and motor systems.

The propensity to use a pointer device to guide visual search merits exploration. It could be that this phenomenon is exaggerated in the aging population, and those with visual impairments. The ability to complete efficient searches, filtering out extraneous noise and distracters has been measured to diminish with age (Owsley and Sloan 1990; Schieber 1994; Orr 1998). It is possible that the use of the mouse pointer or stylus serves to direct attention during visual search.

Major Implication: This suggests that further consideration needs to be made of how visual search changes in the presence of a placeholder or pointing device. The exclusion of the interaction between movement of the mouse, visual search, and the personal abilities that direct both must be considered in modeling interactions. This trend also prompts the need for research to identify if this phenomenon is more prevalent in an aging population and/or for users with limitations to their visual sensory channel. That said, future design recommendation may include the more deliberate incorporation of visual markers to guide search beyond just the stylus or mouser cursor. This may be an additional solution to bridging the performance differential in HCI for aging adults with visual impairments.

Conclusion #3:

> *On the handheld, participants with impaired contrast sensitivity and near visual acuity demonstrated a speed accuracy trade-off in the movement of the card icons*
>
> *.*

Difficulties with the resolution of fine detail and the ability to distinguish a target icon and its features from the background triggered slower rates of performance. In spite of this, these individuals demonstrated a higher level of accuracy in their execution of the task, emergent as a speed accuracy trade-off. The slower rates of visual search and movement time are attributed to the increased obstacle faced in retrieving information from memory that matches what is perceived in the environment.

Supporting Evidence: The eye movement and pupillary response metrics were particularly useful in understanding this effect of speed accuracy trade-off for contrast sensitivity and near visual acuity. Individuals with significantly longer fixation durations and greater deviations in pupillary response also possessed significantly lower levels of contrast sensitivity. Increases in both eye measures are indicative of depth of search and the cognitive load, especially for individuals who experience decrements in contrast sensitivity.

Mean fixation durations that are longer imply more time spent in the interpretation of the visual cue or matching the visual stimulus to an internalized representation. The longer the fixation, the more difficulty the participant has retrieving the internalize representation (Goldberg and Kotval, 1999). This is relevant when considering the effect of decreased contrast sensitivity. Decreases in contrast sensitivity would impose difficulties in perceiving the card from the background, and more importantly could encumber the discrimination of the visual icons and numbers from the background. Therefore, the image they perceive poorly matches what is stored in their internal memory, and fixations are longer as they resolve the differences between these to assure they choose the correct icon. The models of TT and VST decreasing in the presence of improved CS support this.

The group with significantly greater deviation in their pupillary response during the task also had significantly worse CS. The larger the deviation in pupillary response suggests a higher level of mental workload by participants in this group. This lends explanation to the observation of poor CS influencing longer VST, TT, MT, but decreased DD, AD and MDC. These participants exerted a greater amount of effort over the entire course of the task, and in the instance of icon movement to the drop pile, they are more effective, but have to work harder to track the icon movement across the display (e.g. through more fixations, and more mental concentration/workload). The processing of the visual information was quantifiably more involved for those individuals with lower measurable function in contrast sensitivity.

Major Implication: Two major implications emerge from this conclusion. The first is the demonstration of the great utility that eye movement measurement, even summary statistics, can provide in explaining the interaction needs of users. This effect would not have been captured with assessments of the explicit interaction or the participants' exit surveys.

Secondly, this result suggests that contrast should be maximized whenever possible on handheld applications for this population. Like the solution of auditory feedback, improving contrast in a display is also a simple, low-cost solution to improve the interactions of individuals with contrast sensitivity deficiencies, especially those deficiencies associated with aging and/or ocular disease. High contrast environments and sufficient illumination, with the use of the backlight commonly integrated into mobile displays, should be consistently integrated into applications.

Conclusion #4:

Icon spacing triggers longer visual search and movement time, but facilitates a more accurate release of the icon at its destination.

The nature of the effects imposed by increased inter-icon spacing was dependent on the specific component of the task. While the effect of larger spacing was slower performance and longer distances traveled, it also afforded more accurate drop, decreasing the highlight time, and decreasing the probability for over no drop and accidental drop errors.

<u>*Supporting Evidence*</u>*:* In both the clinical and non-clinical models generated on both the desktop and the handheld interaction larger spacing diminished efficiency in visual search, movement time and, in the handheld, drag distance. These results are all highly logical, as increased spacing translated to an increased area of the display over which visual scan was executed, as well as the physical movement of the icon. This contrasts in comparison to the results of Everett and Byrne (2004) and Hornoff (2001). These researchers observed for visually health participants, s a point of diminishing return for changes in icons spacing, in both directions. Apart from the intuitive result that was observed in this dissertation (larger spacing creates a larger area for visual search and movement), these studies demonstrated that decreases in spacing beyond a certain threshold can trigger participants' visual search patterns to be suboptimal (Everett and Byrne 2004) and/or less efficient (Hornoff 2001). This point of diminishing return was not observed in the visual search and selection of the icons in either desktop or handheld of this dissertation research. The participants in this task were not inclined

to change visual search strategies, as the distracter icons were closer to each other. Justification for this discrepancy between the results of this theses with the previous findings are attributable to the difference in age of the participants relative to the much younger, and visually healthy participants who served in the other studies.

Age-related differences have been shown in saccadic eye movements, especially for the acquisition of targets in the peripheral vision. Findings point to a decreased accuracy of saccades for older adults, as more saccades were typically required for this population to fixate on a target in the peripheral vision, or in the context of searching complex visual scenes (see also Kline and Scialfa 1997; Lee, Legge and Oritz 2003). It is highly probably that the older adults are less apt to attend to these items in the periphery, and less likely to experience the negative influence of adjacent icons in visual search and selection.

Icon Spacing was helpful in the drop portion of the task, which suggests the participants could more easily discern when the icon was in correct position to be dropped. The closer the drop piles, the more likely there would be overlap as they distinguish the overlap that is appropriate for the drop of the icon into the pile.

Major Implication: This again demonstrates the importance of the consideration of the entire task when evaluating the effect of different interface manipulations, for improved task validity and generalizability of the results in real world interactions. This result suggests that for older participants, visual search is not compromised in the presence of decreased spacing. However, those tasks require the positioning of an icon over another icon in close proximity to other interface objects are susceptible to negative effects of decreased spacing. An increase in spacing between the objects can afford a more accurate ‘drop.’ Designers need to consider the priorities are for the task, and what the most ‘critical’ interactions are, such as accuracy in drop versus expedience in visual search, and make those design decisions accordingly.

Conclusion # 5:

> *Non-clinically acquired visual function and ocular health measures have potential in characterizing interactions and potential efficacy. Particularly, the VFQ can be highly indicative of an individual’s willingness to accept the assistance offered through visual and non-visual adjustments to the interface.*

The VFQ and VSAT assessments afford an inexpensive means by which to assess visual profile. Although not as comprehensively informative on task interactions across the different task phases as the clinical measures, the VSAT and VFQ showed substantial promise.

Supporting Evidence: Both VSAT and VFQ assessments demonstrated the ability to quantify the impact of changes to the interface in terms of performance. In addition, the VFQ was highly indicative of the participants’ affinity for and willingness to accept changes to the interface, such as larger icons and/or the auditory feedback. In addition, for the handheld, the VFQ tended to be more indicative of interactions involving hand-eye coordination while the VSAT was more indicative of global measures such as visual search and trial time.

Major Implication: While perhaps not the most compelling result to emerge from this research, it critically directs the future of this research. It is the belief that with additional, more sophisticated statistical analyses, that the relationships of VSAT and VFQ will be brought to the surface. Importantly, the results in this dissertation present the potential for relationship. The challenge with these metrics is their susceptibility to account for extraneous of personal and environmental constructs. While both measure aspects of visual function, they also capture an aspect of cognitive functioning and mental health. More sophisticated analyses will be used with these metrics, along with the performance metrics, eye movement data, and exit surveys to better understand these relationships. Additionally, the relationships between the non-clinical and clinical factors merit consideration. It is suspected that the future of these non-clinical measures will be in the validation of clinical measures, and can serve as temporary threshold assessments when participants have not undergone recent clinical ocular examinations.

Conclusion #7:

> *The impact of visual dysfunction and ocular disease demonstrated characteristic differences between the handheld and the desktop PC.*

Visual impairment played a key role the classification of interactions on the handheld and desktop; the most dominant visual factor differed between platforms. The limitations imposed on interaction by visual impairment were clear in both desktop and handheld interaction groups. A universal monotonic trend in diminished performance was observed for the desktop based on declines in near visual acuity. The interaction on handheld was negatively influenced by increases in AMD severity level. The reason for the distinct visual function and ocular health predictors is attributable to the display size relative to the users' visual field.

Supporting Evidence: The difference in the strength of AMD versus NVA is justified between the two different platforms. AMD severity measures the number of drusen on the macula and those with the most severe levels of AMD in an exudative (wet) state. The presence of drusen and patients with wet AMD are more prone to experience the vision loss attributed to the disease. In particular these losses include interruptions, distortions, or loss of fine detail vision in the center of their visual field. The difference between the display sizes is the probably cause of the notable influence of AMD on the handheld, but not in the desktop. That is, the disturbances in the visual field are likely to subtend a larger percentage of the handheld display.

None of the participants, independent of diagnosis, could perceive the entire desktop display at one time (while the specific amount changed). For those participants who worked in the handheld, however, the visually healthy controls could easily account for the entire handheld display in their visual field, while this was not always the case for those with diagnosis of AMD. Individuals diagnosed with AMD, especially at the higher levels of severity would commonly experience an occlusion in their visual field. Between the two platforms, the impact of visual field disruption will effectively occlude a larger percentage of the entire handheld display compared to its imposition on the visibility of the handheld display. Figure 5.3 illustrates this for the handheld and desktop displays. The simulated visual field interruption is the same size, shape, and location on the retina. The percentage of display occluded by the interruption on the handheld is greater than what is occluded by the scotoma in the desktop display.

The effect of increases in set size were subdued, and sometimes absent for participants with AMD, particularly at increased levels of severity. This implies that the participants with more severe levels of disease may not always perceive changes in the display, such as the number of distracters. Unlike the participants with clear uninterrupted vision who may adjust their visual scan strategy at the moment they sense a larger number of icons. Because the entire desktop display is not visible in one visual field, the extent to which the initial scan patterns differ according to disease severity is not as extreme.

Major Implication: Despite the measured affects of AMD severity level on the handheld, it is remarkable that all but one of the participants but one completed all of the trials. These individuals were able to work successfully with the technology, in spite of their disease. Furthermore, the observed gains in performance measured in the interaction terms of AMD and AF provide an easily implemented, low cost solution to mitigate the effects of the disease. Research and commercial efforts both need to address the needs of aging adults in the emergence of new technologies, as they have great potential, and have been demonstrated usable. Furthermore, as shown in Chapter 6, this growing segment of the population, will actively seek the means by which to maintain active, in dependent lifestyles, and will have significant buying power in future markets as the Baby Boomer population ages.

Conclusion #8:

Methodological Contributions: The use of playing card icons maintained the complexity level for the task, but increased the familiarity and comfort for the participants.

The use of playing card as icons in this experiment was a novel approach taken from previous research, which used icons more commonly associated with functions in the Windows desktop computing environment. The card icons were used with much success in this experiment, and were linked to fundamental icon design through a comparison to simple icons, based on their limited details and coding. The relative success of the playing cards in this task warrants their use in future empirical studies concerning icon manipulation.

Supporting Evidence: Based on the background of the participants, the majority of their experience with the computer was not derived from use of word processing, spreadsheet, or other common productivity tools used in the workplace today. Instead, their experience was predictably attributed to 1) games, such as Solitaire or Spider Solitaire, which comes preloaded on Microsoft Windows; and 2) Internet use. The level of familiarity with the 'working icons,' therefore, could not be guaranteed between participants. Based on experimenter observations, the participants were more at ease with the manipulations of these familiar images, and several (with few exceptions) enjoyed their time sorting the playing cards. The use of the playing card eliminated the need to describe, and ensure that participants understood the meaning of the typical Windows Office icons. The use of playing cards provided the means to ensure that the effects measured in this experiment were not attributed to a knowledge and familiarity with icons and their abstractions, but more characteristic of the influence of visual function, ocular disease, personal factors, and interface manipulations.

Major Implication: This study presents results that can serve as a baseline for future studies that utilize the playing card icon in the examination of direct manipulation tasks, independent of icon familiarity and recognition. However, care must be taken in cultural considerations of the participant population, which may direct the level of familiarity with the playing card shapes and numbers. In addition, the playing cards (intentionally) remove the metaphor imposed by the typical Windows icons.

Conclusion #9:

> *Older adults with and without visual impairments are capable of interactions with a handheld PC.*

Another noteworthy result to emerge from this experiment and analyses is the high level of success, comfort, and in some cases, delight that the participants demonstrated in their completion of the playing card task on the handheld. This is the first study of its kind to demonstrate that the usability challenges for this population to use a handheld computer are easily overcome, and that there was not aversion to the device due to its size.

Supporting Evidence: While the handheld employed a only the smallest icon size, 7mmx7mm, *none* of the participants on the handheld task requested to skip trials due to this fact, and *none* of them complained about the icon size (even in the exit interviews). This is curious, when juxtaposed with the number of skipped trials in the desktop that were attributable to icon size, the clear preference by the participants for the largest icon size stated in the exit interviews. The participants, as shown in Chapter 2, were statistically equivalent in terms of their ocular and personal factors between desktop and handheld. In addition, there was a large measured impact of icon size in the models for the desktop. In terms of the display, for those conditions using the smallest icons, the only difference was the amount of white space (or black space in this experiment) on the display and the mouse cursor. The implications of display size should be explored to further clarify what drives the differential in the ability and the perception of capability to complete the task. In our experiment, participants in the desktop setting were more easily overwhelmed in the presence of the larger icon size, to the point where tasks were skipped.

Major Implication: The value of a handheld, mobile computer as a cognitive aid for these individuals, to assist in daily activities to extend their independence has fantastic possibilities. Individuals with visual impairments can make use of a small, handheld, mobile device; with the appropriate design considerations for their abilities.

5.5. Future Work

While the results are not disclosed again in their entirety, the richness of the dataset collected for this dissertation is not overlooked. There are several additional avenues that will be pursued with this data set in the immediate future. In addition, the outcomes of this study stimulate the generation of new research inquiries and initiatives, for the continued investigation of similar interactions, hardware and personal attributes and profiles.

Short term goals, with the present data set

Modeling visual function

Using the various accuracy, efficiency, and information processing measures as the predictor variables, a logistic regression model can be built which estimates the probability for the presence of different ocular health states. Model generation of this type is common in the medical domain, where the risk factors of a medical condition are evaluated in terms of their ability to predict the presence a disease or even its severity level (see Fahrmeir and Tutz 1994 for examples). This is a critical step in the ability to diagnostically adapt a computer interface to personal needs and can provide additional insight on the relationship between the different measures of interaction, not addressed in this dissertation.

Zero-Inflated Poisson regression

The efficacy of VFQ and VSAT, the non-clinical measures of ocular function, were not highly sensitive in the analysis of errors in either the desktop or the handheld. That said, effects of the logistic regression approach, predicting the probability of at least 1 error, should be reconsidered in light of the distributions of the data.

The considerations of different distributions, aside from log should be considered. In several cases, the distributions of errors for these participants appear amenable to the Poisson distribution with the exception of the high frequency with which no errors occurred. Zero-Inflated Poisson Distributions are therefore identified as a potential means for clarifying what is driving the accuracy of the interaction, for both clinical and non-clinical measures, but with particular focus on the extraction of relationships relative to the VSAT and VFQ.

Analytical tools for high frequency data distributions

While the eye movement data provided useful, additional justification for the affects of auditory feedback on the desktop, and contrast sensitivity in the handheld, small nuances in these high frequency pupillary, saccade and fixation data will likely afford the emergence of additional explanations of the outcomes observed in the performance measures. This is especially the case for eye data that come from older participants with ocular disease. These data are highly susceptible to noise external to the information processing mechanisms that are of interest.

Research on a similar data set from participants with AMD in a simple drag and drop task, showed characteristic trends in the high frequency pupillary response signals based on ocular diagnosis with using multi-fractal signal processing techniques (Moloney, Leonard, Shi et al. in press; Shi, Moloney, Pan et al. in press).These trends were not observable with more traditional analytical tools. The exploitation of the high frequency eye movement data can facilitate the extraction of additional explanations of implicit subject differences, their interactions with the task, and ocular condition.

The synthesis of the thresholds informed by the various investigations

The consolidation of the outcomes of the several related studies, including this dissertation studies that contribute to the framework of interaction thresholds is poses a challenge as the result set grows exponentially. The collection of these results can inform designers of the costs and benefits of different design features, in light of user and task related characteristics. However, the synthesis of the study outcomes is less than straightforward. A practical, understandable tool is needed which can demonstrate how to optimize interaction given a set of constraints. Also, one that can be consistently applied under a variety of operating constraints, such as those faced in the design of an interface is necessary since HCI interaction design is dynamically changing in lieu of evolving user, contextual and task related factors.

One avenue that will be explored is the potential for structural modeling, an approached used in Microeconomics, in defining functions, which maximize performance (Henderson and Quandt 1980). Structural modeling applies production functions to evaluate how to optimize to optimize performance at a minimum cost. The combinations of resources considered in production functions, both fixed and variable, are evaluated to optimize the best combination to a. A production process is highly analogous to the HCI, as both are time based and work based on available resources and constraints. The process is the computer interaction, the variable inputs are the options for the interface design features (e.g., size, spacing auditory cues, etc.), and the fixed inputs in the function are captured by personal factors of the user, including their visual profile. The single output of the function is the outcome measure, defined by the goal of the interaction, and informed by the framework of performance thresholds (e.g., moving the card to the pile as quickly and accurately as possible). The optimization of this function could inform designers of most appropriate set of design features are incorporated to optimize interaction for the specified participant profile

Long-term research agenda

The inclusion of contextual constraints and situationally-induced impairment in investigations of interactions for users with visual impairments

Based on the high levels of task successes that are consistently observed for the participants in this, and other studies with visual impairment and HCI The affects of time limits or other contextual constraints such as task sharing, is an important next step a more complete characterization and development of interaction thresholds. A research unifying work in both Situationally Induced Impairments (SII) and Disability Induced Impairments (DII) is a notable milestone, and particularly critical step for the successful proliferation of mobile devices to users with a variety knowledge and range of capabilities.

Continued exploration of interactions with small mobile displays

Based on the positive feedback and successful interactions with the handheld in this sampled user population, the development of applications for handheld, mobile computers that can extend the independence of older adults should be a prime research initiative. Coupled with the development of applications and devices is the continued empirical investigation of the interactions required on the handheld, as it was observed to focus on different resources than tradition, desktop computing. Furthermore, these empirical investigations can generate practical design guidelines that inform more inclusive technologies through simple design choices, such as increases in contrast.

Ongoing framework development

Empirically driven examinations of the critical interactions are needed for the continued expansion of the framework of interaction thresholds, as they relate to visual profiles. Continued, controlled empirical research investigations similar what has been presented in this dissertation are necessary to sufficiently anticipate the interaction needs of individuals as new interaction paradigms introduced by emergent information technologies. A comprehensive understanding of the interaction and the user abilities facilitate design affords interaction for as many users as possible. This serves to bridge, as oppose to widen, the digital divide for users with divergent needs for uninterrupted access to information technology.

5.6 REFERENCES

Byrne, M. D. (1993). Using icons to find documents: Simplicity is critical. Proceedings of the SIGCHI conference on Human factors in computing systems, Amsterdam, The Netherlands.

Edwards, P. J., L. Barnard, V. K. Emery, J. S. Yi, K. P. Moloney, T. Kongnakorn, J. A. Jacko and F. Sainfort (in press). "Understanding users with Diabetic Retinopathy: Factors that affect performance in a menu selection task." Behaviour & Information Technology.

Edwards, P. J., L. Barnard, V. K. Emery, J. S. Yi, K. P. Moloney, T. Kongnakorn, J. A. Jacko, F. Sainfort, P. R. Oliver, J. Pizzimenti, A. Bade, G. Fecho and J. Shallo-Hoffmann (2004). Strategic Design for Users with Diabetic Retinopathy:Factors Influencing Performance in a Menu-Selection Task. The Sixth International ACM SIGCAPH Conference on Assistive Technologies (ASSETS 2004), Atlanta.

Emery, V. K., J. A. Jacko, T. Kongnakorn, V. Kuruchittham, S. Landry, G. M. N. III, A. Sears and J. Whittle (2001). Identifying critical interaction scenarios for innovative user modeling. 10th International Conference on Human-Computer Interaction, New Orleans, LA.

Everett, S. P. and M. D. Byrne (2004). "Unintended effects: Varying icon spacing changes users' visual search strategy." CHI Letters **6**(1): 695-702.

Fahrmeir, L. and G. Tutz (1994). Multivariate Statistical Modelling Based on Generalized Linear Models. New York, Springer-Verlag.

Henderson, J. M. and R. E. Quandt (1980). Microeconomic Theory: A Mathematical Approach. New York, McGraw-Hill.

Hornoff, A. (2001). "Visual search and mouse-pointing in two-dimensional visual hierarchies." ACM Transactions on Computer-Human Interaction **8**(3): 171-197.

Jacko, J. A. (1999). "Incorporating clinical diagnoses in the prediction of performance on computer-based tasks for low vision users." ACM SIGCAPH Computers and the Physically Handicapped **64**(June 1999): 3-5.

Jacko, J. A. (2000). "Visual impairment: The use of visual profiles in evaluations of icon use in computer-based tasks." International Journal of Human-Computer Interaction **12**(1): 151-164.

Jacko, J. A. (2000). "Visual impairment: The use of visual profiles in evaluations of icon use in computer-based tasks." International Journal of Human-Computer Interaction. **12**(1): 151-164.

Jacko, J. A., L. Barnard, T. Kongnakorn, K. P. Moloney, P. J. Edwards, V. K. Emery and F. Sainfort (2004). "Isolating the effects of visual impairment: Exploring the effect of AMD on the utility of multimodal feedback." CHI Letters **6**(1): 311-318.

Jacko, J. A., A. B. Barreto, J. Y. M. Chu, H. S. Bautsch, G. J. Marmet, I. U. Scott and R. H. Rosa (2000). Using eye tracking to investigate graphical elements for fully sighted and low vision users. ACM conference on Eye Tracking Research Applications (ETRA), Palm Beach, FL.

Jacko, J. A., A. B. Barreto, G. J. Marmet, J. Y. M. Chu, H. S. Bautsch, I. U. Scott and R. H. Rosa (2000). Low vision: The role of visual acuity in efficiency of cursor movement. ACM Conference on Assistive Technologies, Arlington, Virginia.

Jacko, J. A., A. B. Barreto, I. U. Scott, J. Y. M. Chu, H. S. Vitense, F. T. Conway and W. B. Fain (2002). "Macular degeneration and visual icon use: Deriving guidelines for improved access." Universal Access and the Information Society **1**: 197-296.

Jacko, J. A., K. P. Moloney, T. Kongnakorn, L. Barnard, P. J. Edwards, V. K. Emery, F. Sainfort and I. U. Scott (in press). "Multimodal feedback as a solution to ocular disease-based user performance decrements in the absence of functional visual loss." International Journal of Human-Computer Interaction.

Jacko, J. A., K. P. Moloney, T. Kongnakorn, L. Barnard, P. J. Edwards, V. K. Leonard, F. Sainfort and I. U. Scott (2005). "Multimodal feedback as a solution to ocular disease-based user performance decrements in the absence of functional visual loss." International Journal of Human-Computer Interaction **18**(1): 183-218.

Jacko, J. A., R. H. Rosa, I. U. Scott, C. J. Pappas and M. A. Dixon (1999). Linking visual capabilities of partially sighted computer users to psychomotor task performance. 9th International Conference on Human-Computer Interaction, Munich, Germany.

Jacko, J. A., I. U. Scott, A. B. Barreto, H. S. Bautsch, J. Y. M. Chu and W. B. Fain (2001). Iconic visual search strategies: A comparison of computer users with AMD versus computer users with normal vision. 10th International Conference on Human-Computer Interaction, New Orleans, LA.

Jacko, J. A., I. U. Scott, F. Sainfort, L. Barnard, P. J. Edwards, V. K. Emery, T. Kongnakorn, K. P. Moloney and B. Zorich (2003). "Older adults and visual impairment: What do exposure times and accuracy tell us about performance gains associated with multimodal feedback?" CHI Letters **5**(1): 33-40.

Jacko, J. A., I. U. Scott, F. Sainfort, K. P. Moloney, T. Kongnakorn, B. Zorich and V. K. Emery (2002). Effects of multimodal feedback on the performance of older adults with normal and impaired vision. Lecture notes in computer science. C. S. Noelle Carbonell. NEED CITY, Springer. **2615:** 3-22.

Jacko, J. A. and A. Sears (1998). Designing interfaces for an overlooked user group: Considering the visual profiles of partially sighted users. ACM Conference on Assistive Technologies, Marina del Rey, California.

Kline, D. W. and C. T. Scialfa (1997). Sensory and perceptual functioning: Basic research and human factors implications. Handbook of Human Factors and the Older Adults. A. D. Fisk and W. A. Rodgers. New York, Academic Press**:** 28-54.

Kline, R. L. and E. P. Glinert (1995). Improving GUI accessibility for people with low vision. SIGCHI Conference on Human Factors in Computing Systems, Denver, Colorado.

Lee, H.-W., G. E. Legge and A. Oritz (2003). "Is word recognition different in central and peripheral vision?" Vision Research **43**: 2837-2846.

Macular Degeneration Partnership. "What is AMD?" Retrieved November 5, 2002, from http://www.macd.net/amd/index.asp.

Moloney, K. P., V. K. Leonard, B. Shi, B. Vidakovic, J. A. Jacko and F. Sainfort (in press). "Leveraging data complexity: New mehtods for analyzing pupillary behavior of older adults with viusal imparimetn during performance of psychomoter tasks." Transactions on Computing and Human Interaction.

Orr, A. L. (1998). Issues in Aging and Vision. New York, American Foundation for the Blind.

Owsley, C. and M. E. Sloan (1990). Vision and aging. Handbook of neuropsychology. F. Boller and J. Grafman. Amsterdam, Elsevier. **4:** 229-249.

Partnership, M. D. (2005). "Anatomy of the Eye." Retrieved May 5, 2003, from http://amd.convio.net/site/PageServer?pagename=Anatomy_of_the_Eye.

Schieber, F. (1994). Recent developments in vision, aging and driving: 1988-1994, Transportation Research Board.

Scott, I. U., W. J. Feuer and J. A. Jacko (2002). "Impact of graphic user interface screen features on computer task accuracy and speed in a cohort of patients with Age-related Macular Degeneration." American Journal of Ophthalmology **134**(6): 857-862.

Scott, I. U., W. J. Feuer and J. A. Jacko (2002). "Impact of visual function on computer task accuracy and reaction time in a cohort of patients with Age-related Macular Degeneration." American Journal of Ophthalmology **133**(3): 850-857.

Shi, B., K. P. Moloney, Y. Pan, V. K. Leonard, B. Vidakovic, J. A. Jacko and F. Sainfort (in press). "Classification of High Frequency Pupillary Response using Schur Monotone Descriptors in Multiscale Domains." Journal of Statistical Computing and Simulations.

Tun, P. A. and A. Wingfield (1997). Language and communication: funndmentals of speech communication and language processing in old age. Handbook of Human Factors and the Older Adult. A. D. Fisk and W. A. Rodgers. New York, Academic Press**:** 125-149.

Vitense, H. S., J. A. Jacko and V. K. Emery (2002). "Foundation for improved interaction by individuals with visual impairments through multimodal feedback." Universal Access and the Information Society **2002**(2): 76-87.

Vitense, H. S., J. A. Jacko and V. K. Emery (2003). "Multimodal feedback: an assessment of performance and mental workload." Ergonomics **46**(1-3): 68-87.

CHAPTER 6

EXPLORING THE COMMERCIAL POTENTIAL OF ASSISTIVE TECHNOLOGY RESEARCH

6.1. <u>Introduction</u>

The applications of the research presented in this dissertation, and other associated studies emergent from the Laboratory for Human Computer Interaction and Health Care Informatics (HCI) at Georgia Tech, are not limited to basic research and academe. The results, principles and theories evolving from this work can be applied to commercial technologies. To this end, the research feeding this dissertation was the focus of work in the TI:GER program (Technological Innovation: Generating Economic Results), a multidisciplinary program in the College of Management at Georgia Tech, under the leadership of Dr. Marie Thursby.

The TI:GER Program, sponsored by a NSF IGERT grant, and directed by Dr. Marie Thursby, matches PhD research projects with MBA students from Georgia Tech and JD students from the Emory School of Law. These multidisciplinary teams explore the intellectual property landscape pertaining to the projects, and posit the most promising opportunities for the eventual application of the research. As a two-year program, the final deliverable is a business plan. The author of this dissertation was accepted into the TI:GER program in the Fall of 2002, and worked for two years with her teams to understand and develop an executable business plan for the research.

Prior to the TI:GER program, the application of this specific research had been considered, sans the specifics of the business plan (e.g., path to market, funding, exit strategy). A grant was awarded to Dr. Julie A. Jacko from the Intel Research Corporation to explore the concept of adaptive, multimodal software that could automatically change the computer interface for individuals, contingent on their visual capacity. Research studies sponsored through this award generated an identification of the interactions vital for GUI interaction, thresholds for performance levels for adaptive changes under varied degrees of visual function, as well as the efficacy of multimodal feedback for the most critical GUI interactions. Finally, a 'proof of concept'/beta version of the interface was produced.

While the underlying empirical research and the beta version were under development in the HCI lab, work within the TI:GER program explored the practicability for the application of the research, the intellectual property, as well as potential future applications. The business and commercialization plans were iterated several times over the course of the TI:GER program, and the project was a successful competitor in a handful of invited commercialization and business plan competitions:

- Finalist, University of San Diego Business Plan Competition San Diego, California, April 2004
- 3rd Place, Georgia Tech Business Plan Competitions, Georgia Institute of Technology, Atlanta, Georgia, March 2004

- 1st place Faculty Advisor's Award, I2P Commercialization Plan Competition, University of Texas, Austin, Texas, November, 2003
- 3rd place Award, I2P Commercialization Plan Competition, University of Texas, Austin, Texas, November, 2003

After participation in the TI:GER program came to an end, this work maintained interest. The business plan and product concept continued to evolve based on positive, constructive feedback provided during the competitions. The most recent instantiation of this plan was presented by David Beck and Michael Orndorff, two Georgia Tech MBA alumni ('05), at the 2005 MOOT Corp International Business Plan Competition held in Austin, Texas. The plan did place runner up to the finalists, and received a great deal of positive feedback from judges and venture capitalists attending the event.

This chapter presents the most recent version of this business plan. This plan represents a culmination of the collaborative effort between V. Kathlene Leonard, the MBA students and JD students over the course of three years, working under the advisement of Dr. Thursby and Dr. Jacko. At the time of the completion of this dissertation, no formal plans have been made to follow-through with this plan, but opportunities are continually explored. Nevertheless, this plan demonstrates the practical applicability of this type of research, and the extent of the impact and success it could eventually engage in the market place. This is exceptionally valuable in terms of the assistive technology domain. The assistive technology community is continually imposed upon by barriers to market for solutions that have the potential to be life-altering for the target consumers.

There is heightened potential for accessible technologies to proliferate into mainstream consumer marketplace with the support of a team with interdisciplinary expertise in marketing, finance, management and intellectual property. Together these skill sets to can cultivate and foster solutions aimed at this market segment and ensure a sustainable position in the market. The following business plan makes evident the growing, critical need for technologies to extend independence, such as the solutions informed by this dissertation. The business plan in serves as a model for the extension of selective components of assistive technology research into the mainstream. This is the only known documentation that details an example of the resources and steps necessary to translate a concept in assistive technology and universal design research from the lab and into a practical application and use in the marketplace, with great promise for profitable returns and sustainability.

6.2. Crossing Point Technologies Business Plan, April 2005

Executive Summary

Today there are over 10 million Americans who are unable to effectively use computers due to visual impairments. This number is set to rise sharply as a result of the aging Baby Boom population.

Crossing Point Technologies enhances the accessibility of technology for people with visual impairments. Integrating customizable accessibility software solutions through the use of innovative research, Crossing Point enhances individuals' abilities to interact with computers and other graphical user interfaces (GUIs). Unlike any device currently on the market, this technology will increase productivity and access to information technology through intelligent, customizable solutions that improve customers' quality of life.

Background

Crossing Point provides universal accessibility to GUI's for users with visual impairments. These are people with vision problems that are uncorrectable with glasses or surgery, and they include individuals with impairments such as macular degeneration and diabetic retinopathy. The significance of this section of the population is notable, as is their need for accessibility. According to Forrester Research, over

> I am excited about the solution Crossing Point has developed and hope to be one of the first customers of the product!
>
> -Dave Ostrowski
> Product Manager, ISS
> Visually Impaired

33 million computer users of working age have visual impairments that hinder their abilities to use computers. Of this number, over 10 million sighted individuals have a *severe need* for accessibility technology based on the severity of their impairment. Crossing Point estimates the market for these individuals to be approximately $300 million a year. Additionally, this market is growing at a rapid pace. In the next 10 years, there will be 2.5 times more adults aged 65 to 74 years using computers as there are today.

Problem

People with visual impairments have trouble using computers, and current interfaces tend to be inflexible in the ways in which they interact with users. With current solutions, the user bears the responsibility of determining how to improve the interface, either by changing something about themselves (e.g. getting reading glasses), or by changing system settings (e.g. Microsoft Windows accessibility options). Crossing Point's products transfer the burden and knowledge requirements of identifying and implementing assistive changes from the user to the computer.

Crossing Point's Solution: EyeAbility

Based on over $1 million in research spanning five years, Crossing Point's first product, EyeAbility, creates an optimum interface experience for people who are visually impaired. EyeAbility is a software solution that allows the computer to adapt to the user's unique abilities and impairments rather than forcing a user with impairments to attempt to adapt to an interface that was designed for people with normal vision. EyeAbility contains three core components. First, the user completes a performance assessment, which enables the system to build a profile of the individual's unique abilities and impairments. Second, an automated diagnostic tool, the core of Crossing Point's intellectual property, takes the performance results and creates a usability intervention set, a profile of the optimum interface for the specific user. Finally, based on the results of the diagnostic, the third component recommends the appropriate interface interventions and, with the user's permission, affects the recommended changes.

EyeAbility is differentiated and has a competitive advantage over other products because it is based on years of research correlating a user's performance on specific tasks to the appropriate interventions across the spectrum of modalities needed to optimize the computer interface for that specific user. The resulting software is *comprehensive, customized*, and *adaptive*.

Management Team

Crossing Point's executive team brings a wide range of relevant skills and startup experience to the company. Together, they bring experience starting, growing, and successfully exiting investor-funded ventures. They have experience in Human Computer Interaction (HCI) technology development, and they have expertise in bringing complex new products to market. Crossing Point's founding executives are:

David Beck, CEO – With thirteen years of leadership experience in operations management and technology commercialization, Beck has opened several new ventures and consulted with other entrepreneurs on successfully starting and growing their businesses.

Michael Orndorff, COO – The former CEO and founder of MultiMediums, Inc., an application development firm acquired by MarketingCentral in 2002, Orndorff has experience in software product development and in bringing new products to market.

V. Kathlene Leonard, CSO – The inventor of Crossing Point's technology and leader of the company's ongoing research efforts, Leonard is an expert in the human aspects of personal and networked computing, as well as universal access to information technologies.

Crossing Point has also enlisted the aid of a strong board of advisors, which includes the director of Georgia's largest vision center (Subie Green), an expert on HCI research and development (Dr. Julie Jacko), the leader of a technology commercialization lab (Steve Derezinski), a direct and relationship marketing executive (Ann Bachrach), and the manager of a nationally recognized business incubator (Tony Antoniades). In addition, Crossing Point has secured the legal services of Nelson Mullins, a law firm with specific expertise servicing startup companies, as well as the business advisory services of Gross Collins, a local accounting and consulting firm.

Company Status

Crossing Point started operations in January 2005 and became a member company of VentureLab company (commercialization center for the state of Georgia) in March 2005. The company currently has 3 employees and is finalizing seed stage funding in the amount of $260,000. Also, Crossing Point (under the previous name InfoVision) recently became the subject of a Harvard Business School case about navigating the university technology transfer process. The case will be published in the coming months, and Harvard has plans to follow the company through the commercialization process.

Financials

Crossing Point forecasts break even in 2007, with total revenues of $13 million by 2009, as shown in Table 6.1. To reach this stage, Crossing Point will need $3,000,000 in Series A funding in 2006. The company is looking for investors who will be able to provide not only capital, but also the experience and industry relationships necessary to help the growth of the business.

Table 6.1. Forecast financials.

(In thousands)	**2005**	**2006**	**2007**	**2008**	**2009**
Sales	$0	$782	$2,939	$6,726	$13,043
EBIT	($187)	($807)	$661	$4,104	$8,865
Net Income	($187)	($807)	$661	$2,667	$5,762
Operating Cash Flows	($187)	($1,181)	$32	$1,488	$3,735

Company Overview

Formed in January 2005, Crossing Point is an assistive technology software company focused on meeting the needs of a large and growing population of computer users with visual impairments. The ideas behind Crossing Point Technologies were originally conceived in the work of Dr. Julie Jacko and further developed in collaboration with CSO, V. Kathlene Leonard and Georgia Institute of Technology (Georgia Tech). This collaboration has progressed over the past 4 years from their research in the Laboratory for Human Computer Interaction and Health Care Informatics associated with Georgia Tech in Atlanta, Georgia. Dr. Jacko is a leader in the field of human computer interaction (HCI). She works extensively in the areas of accessibility and universal design, specifically for individuals with visual impairments. The underlying concepts of EyeAbility emerged from empirical research conducted with individuals with age-related macular degeneration and diabetic retinopathy. Dr. Jacko has received over $1 million in grant funding from the National Science Foundation and Intel Research Corporation for this work.

In March 2005, Crossing Point was accepted into VentureLab. VentureLab, part of the Office of Economic Development and Technology Ventures, offers assistance throughout the commercialization process, helping evaluate the commercial value of an innovation, connecting faculty with entrepreneurs who have the track record necessary to attract outside funding, and offering pre-seed awards to help move innovations to the commercial stage. Crossing Point is currently working with VentureLab to secure a $50,000 Phase 1 Georgia Research Alliance Commercialization Grant.

Product

> From a productivity standpoint, this product would definitely be attractive.
>
> - David Brookmire, Owner, Corporate Performance Strategies HR Consulting Firm

In the last twenty years, graphical user interfaces (GUIs) have fundamentally changed the way people use computers. Users are able to more easily learn and use computers by operating within the context of a metaphor (e.g. the well known ‘office' metaphor employed by Microsoft Windows). However, by placing such a high reliance on the visual sensory channel, individuals with visual impairments find significant barriers to effectively using computers. These visual impairments may hinder the use of technical applications such as email, file organization, spreadsheets, and other business applications. Crossing Point’s technology creates an optimum interface experience for people who are visually impaired.

With these goals in mind, there are several key features of the solution that enable Crossing Point to deliver the desired benefits:

Customizable solution: The solution is dynamically tuned to deliver the appropriate combination for the individual, taking into account their unique abilities and impairments.

Research based metrics: The solution utilizes a proprietary database of thousands of performance threshold metrics collected from directed research.

Multimodal, adaptive interventions: Based on years of research, a mixture of adaptive interventions can be deployed based on the needs of the user. These interventions do not solely rely on traditional visual-based GUI elements, rather the solution includes the ability to integrate tactile, auditory, and additional visual input and output mechanisms.

Portable accessibility: Developing a better interface is only half the battle – there must also be a way for the user to have access to the improved interface from multiple locations. This enables a user to work with the same interface enhancements from work, home, and any other computer. This same feature also allows multiple users to have separate interface profiles on the same machine, allowing each to have their own customized, optimized solution.

Expandability: The existing and future research that forms the basis of Crossing Point's products applies to displays beyond traditional displays. Initial research has already proven the applicability of these principals to handheld devices, and additional research will allow Crossing Point to move this platform technology to such interfaces as ATMs, point of sale kiosks, PDAs, cell phones, and more.

Product Vision

The primary software platform developed by Crossing Point Technologies will contain three core components. The three components work together to systematically deliver the benefits in a compelling, cost-effective manner: 1)

Performance Assessment; 2) Automated Diagnostic; and 3) Multimodal Interventions.

Performance Assessment

The performance assessment component instructs the user to perform a series of computer tasks. These series of tasks enable the system to build a profile of the individual's unique abilities and impairments with regards to computer usage.

While the user is performing the requested tasks, the system is logging multiple aspects of the user's actions, including the amount of time it takes to perform each task, how many incorrect attempts are made, the screen location in which the task is being performed, etc. This data forms the basis of the capabilities profile that is later analyzed by the automated diagnostic component.

The tasks that the user performs represent the key usability components of today's modern GUIs, such as operating file menus, performing drag and drop, performing text entry, and highlighting text. These GUI interactions form the basis for interacting with most applications on the market today. By testing the individual's abilities with regards to these interactions, the system is able to form a complete picture of his/her abilities with regards to the computer as a whole.

The performance assessment dynamically adjusts subsequent tasks, spending greater time in areas determined to be problematic and skipping over in-depth testing of areas in which the user is not having problems. In this way, the performance-testing period can build a complete profile of the user's computer interactions while minimizing the time it takes the user to complete these tasks. The complete performance assessment regimen will take between 5 and 30 minutes, depending on the impairment level and the degree to which the user is familiar with computers.

Automated Diagnostic

Once the task performance profile has been created for the user, this data is sent via the Internet to the automated diagnostic component. It is in this component that the real strength of Crossing Point's offering resides. This component takes the newly created performance results and creates a usability intervention set, a series of recommended changes and utility installations, in order to optimize the user's computer interface.

This is done by comparing the user's performance data to thousands of interaction performance threshold metrics from users across a range of visual abilities in similar tasks. The threshold metrics rely upon values aggregated from extensive research, carried out by the Crossing Point researchers. This threshold data is absolutely essential to the process and represents the key to Crossing Point's competitive advantage. Based on years of research, Crossing Point holds a proprietary database of knowledge of how people with visual impairments interact with computers through a host of performance measures and vision assessment metrics.

Based on a user's interactions with each diagnostic task, the system automatically adapts itself to optimize each user's interactions. The automated changes will be made to the interface with respect to visual, auditory, and haptic interaction modalities. The proprietary knowledge of real interaction by people with visual impairments drives this intelligent system so that the software is customizable.

The research shows that users with different visual profiles benefit from different combinations of feedback, and that all users benefit from some level of augmentative feedback. By performing and analyzing the appropriate diagnostic tests, the diagnostic component can then assign an appropriate intervention profile, which stores the inclusive results of their diagnoses and recommended adaptations. Screenshots of the diagnostic in the Alpha version of EyeAbility are located at the end of this plan.

Interface Interventions

This is the final component of the software and is illustrated in Figure 6.1. It is in this component that the results from the previous two components are made usable. The results of the diagnostic are used to recommend a series of changes to the interface customized for the individual user. Upon accepting the recommendations, the Crossing Point product will automatically put the changes into action. While advanced users will have the option to accept or dismiss each individual intervention, the typical user will be able to have the system optimize their entire system in one smooth action.

First, the system recommends and implements changes to the underlying operating system that will enhance the computer's usability for that specific individual. These changes include optimizing settings in Microsoft's Accessibility Options, enabling appropriate Microsoft Accessibility Tools, and applying desktop themes that have been determined to increase usability for users with similar abilities.

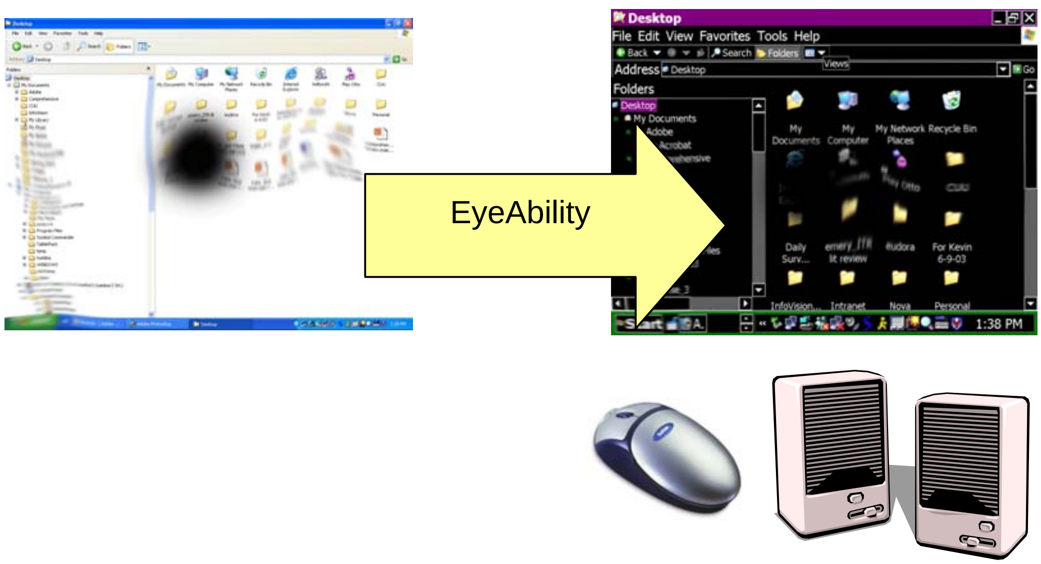


Figure 6.1. Illustration of the EyeAbility concept, how it changes the interfaces for an individual with a central visual field deficit, such as that common macular degeneration.

Next, additional utilities, some developed by Crossing Point as well as others offered as freeware, will be recommended to address impairments not effectively covered by the tools built into the operating system. These applications will be installed, on an as-needed basis, at this point.

Additionally, the user's computer is searched for specific business applications for which Crossing Point has custom application profiles. If any of these applications are found to be installed, Crossing Point can directly affect changes to the application's interface by using build-in Application Programming Interfaces (API's). For example, custom interface settings can be changed for applications such as CRM, SFA, and other business applications.

Finally, in the case where a user's intervention profile indicates that implementation of the preceding interventions is not enough, a series of third-party software applications will be recommended that can increase usability for the user even more. This includes accessibility applications such as JAWS, Magic, and Zoomtext that are targeted at users who have more severe impairments. Ultimately, purchasing and installation capabilities will be integrated into the product. This will allow a user, after receiving a recommendation to use a third party product, to automatically purchase, download, install, and configure the third party product directly from the Crossing Point application.

Initial Offering

The initial product will be focused specifically on users with vision impairments. Ultimately, the product will be targeted to assist users with a broader range of problems, including manual dexterity and hearing impairments. Crossing Point is focusing on the low vision market for two reasons: individuals with low vision disabilities represent the largest sub-market and all the research to date has focused on this area.

This initial product, EyeAbility, will be released in three editions – Home, Professional, and Enterprise. Each edition offers the core EyeAbility platform focused for a different market demographic.

EyeAbility Home Edition

As the name implies, this edition of the product is targeted for use by home users. This version contains the three core components; however, this edition does not include the business application intervention profiles nor does it include support for universal profiles. Instead, the application and resulting intervention is licensed for use by one user on one machine. Additionally, it can only be installed on computers for use in non-commercial or educational settings.

EyeAbility Professional Edition

The Professional Edition builds on the capabilities of the home edition and includes the business application profiles in the intervention component. This edition can be installed on machines being used for business purposes. Additionally, this license allows for an individual to use his/her profile on any machine. This enables Professional users to use the improved interface on their home machine, as well as any other machines they wish. It is important to note that given the customized nature of the resulting interface, the results are only valuable to the user who performed the test and diagnostic components.

EyeAbility Enterprise Edition

This edition includes the same features as EyeAbility Professional, but is available in a site license format. This enables a company to purchase the rights for any employees within their organization to use the product without the need for additional licenses. As with the Professional Edition, this edition also gives users the ability to use it from home as well.

Proprietary Position

The initial research that forms the basis for Crossing Point's technology was performed by Dr. Julie Jacko and Kathlene Leonard at Georgia Tech. The initial database and software has been licensed by Crossing Point from the Office of Technology Licensing (OTL) at Georgia Tech. This license grants full, exclusive, assignable rights to all the performance data as well as all software developed during their research (including the testing software and the resulting diagnostic software). To keep the company's capital structure clean, OTL's proposed licensing terms stipulate a one-time cash payment upon an exit event of 1% of the purchase price.

The Market

Market Overview

According to Forrester Research, Inc., in 2004 there were over 130 million computer users of working age (between 18 and 64 years old). Of this population there are 33 million people with visual impairments who are either likely or very likely to benefit from the use of accessible technology, illustrated in Figure 6.2 and Table 6.2.

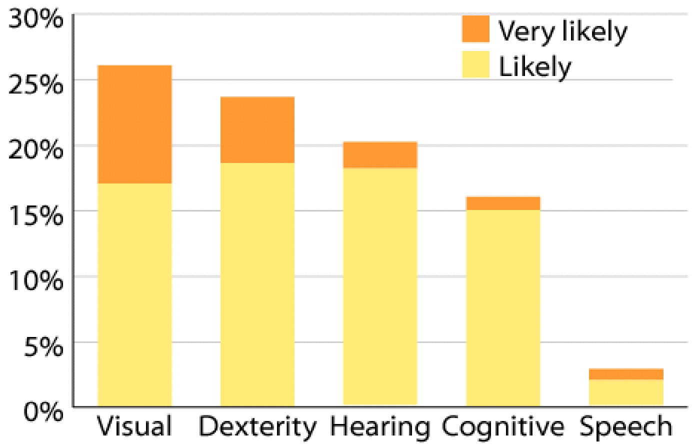


Figure 6.2. Percentage of working age (18-64 years) Americans who are either likely or very likely to benefit from different classes of accessible technology interventions.

Table 6.2. Number of working age (18-64 years) Americans who are either likely or very likely to benefits from different classes of accessible technology interventions.

	Likely (millions)	**Very likely** (millions)	**Total**
Visual	21.9	11.1	**33.0**
Dexterity	24.4	6.8	**31.2**
Hearing	24.0	2.5	**26.5**
Cognitive	19.5	1.7	**21.2**
Speech	2.5	1.1	**3.6**

Crossing Point will initially focus efforts on over 11 million people with visual impairments who are categorized as very likely to benefit from accessible technology. These are individuals who have self-identified as having problems interfacing with their computers. According to the study, this group includes:

Individuals who reported having an impairment that limits employment.

Individuals who reported difficulty with all of the tasks within a difficulty/impairment type some of the time and report having an impairment.

Individuals who reported difficulty with most of the tasks within a difficulty/impairment type most of the time.

Examples of severe visual difficulties and impairments include having non-correctable visual problems that cause difficulty performing many visual-related tasks. Specific problems include macular degeneration, diabetic retinopathy, and retinitis pigmentosa. This group also includes approximately 1 million individuals who are blind and could not benefit from Crossing Point's software; therefore, the addressable target market of working age computer users with visual impairments is just over 10 million people.

Market Growth Opportunities

Aging Population

The anticipated surge of the baby boom population (born 1946-1964) into the classification of older adults generates a new group of consumers with identifiable needs. A large part of this group uses computers on a regular basis. Unless technologies provide such users with features that help them through their continued visual loss, they will lose a large part of their independence. Crossing Point's ability to adapt an interface to the needs of its users and to anticipate users' needs will be beneficial to older adults with visual impairments.

During the next 15 years, over 82 million people who constitute the baby boom population will join the older adult population (Census 2000 Brief), as illustrated in Figure 6.3. And, in 10 years, 2.5 times more adults aged 65 to 74 years will be using computers as there are today. This is a new trend for this generation. As baby boomers enter retirement, they will continue to use computers actively and will demand that the technology work around the difficulties/impairments that they will experience as they age. The growth in computer use by this demographic will play a particularly pivotal role in extending the reach of accessible technology.

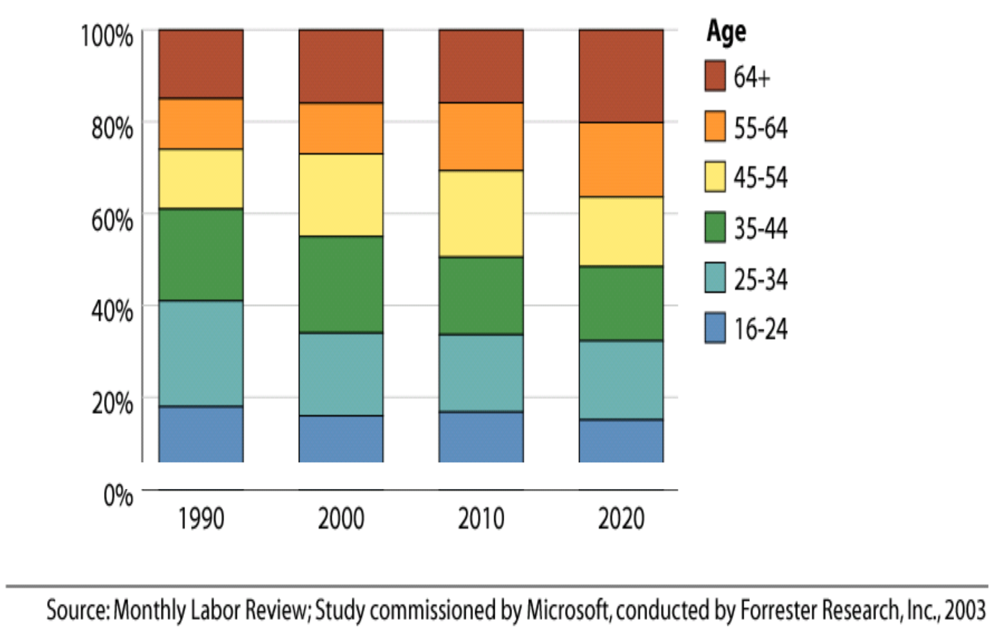


Figure 6.3. Past, current, and future projections of the distribution of age groups within the US populaion.

Older adults, aged 65 and older, are known to experience age-related changes in memory, sensory perception, and other aspects of cognitive and motor processes. These differences between the older adult population and younger populations warrant specific consideration in the design of computer systems to be inclusive of older adults in their user base. Figure 6.4 details how difficulties and impairments increase with age.

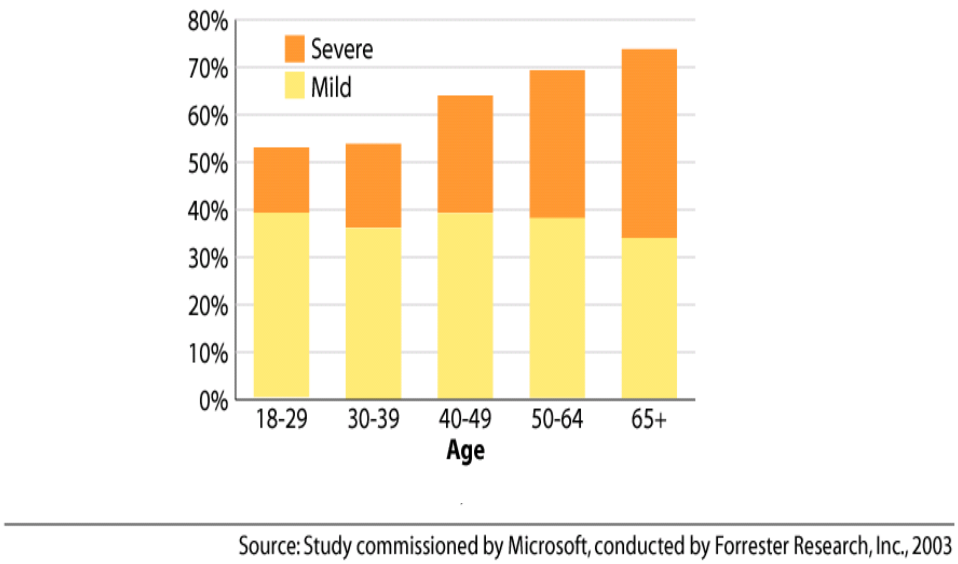


Figure 6.4. Percentage of Americans with either mild or severe impairment by age group.

Baby boomers have incorporated computers with GUI's into their work and personal lives, becoming well-connected people. As this population ages, its members will want to sustain their use of technologies, especially since such technology use offers a means to maintain independence. Presently, consumers who are older than 65 are not as technically savvy as those in younger age segments. The aging of the baby boomer population soon will create an older consumer group with a demand for technology. The marketplace has to anticipate this growing need and the unique requirements of older users when creating computer technologies.

Computer experience is one trait that clearly sets the future elderly population apart from the current elderly population. According to the 2000 US Census, only 28% of adults aged 65 and older have home computer access compared to 51% for adults aged 55-64 and 65% for those aged 45-54. As the baby boom population ages, it will be the first generation in which the majority of the members will already have significant computer experience when they reach the age of 65. A sufficient understanding of the influence that computer experience has on older adults' interactions with computers and applications can enable designers to anticipate interface needs for this evolving user group.

Future Opportunities

Going forward, Crossing Point believes that there are compelling opportunities for market growth by expanding applicability on two fronts.

Addressing other functional impairments. One example is individuals with manual dexterity issues. They comprise the second largest group to report a need for accessible technology, and, particularly at the older end of the age spectrum, there is a definite overlap of individuals with both visual and manual dexterity impairments.

Expanding use of accessible technology to a wider audience of computer users. Crossing Point's technology can be used to address customers with a lower level of impairment. These are individuals whose needs are not as pressing as the initial market, but who could benefit from use of Crossing Point's products. The overall growth of accessible technology users in the coming years is predicted to be substantial, as shown in Figure 6. 5.

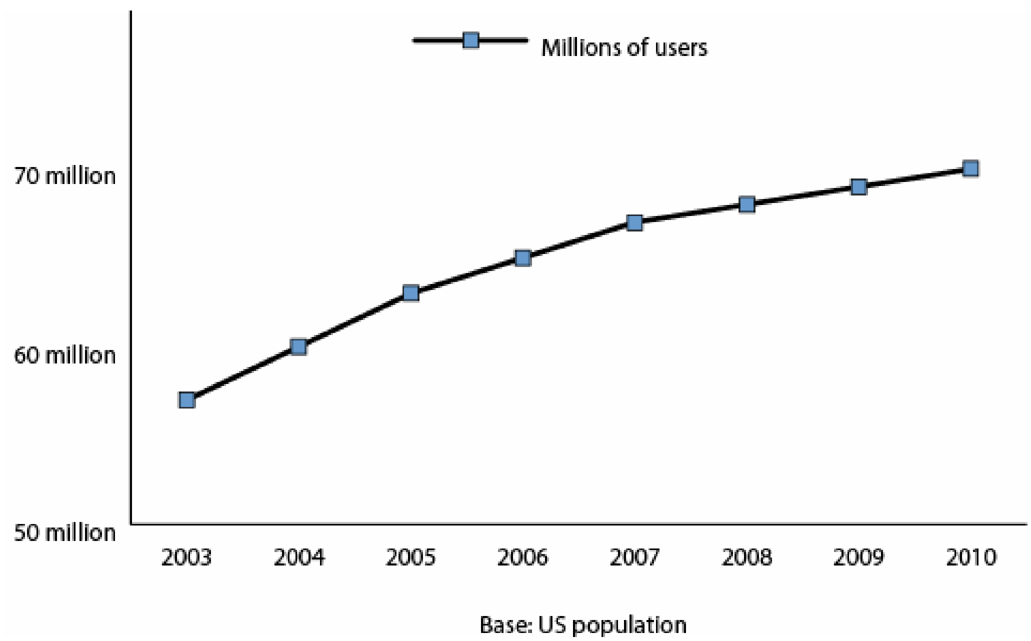


Figure 6.5. Predicted growth in number of accessible technology users from 2003- 2010.

Competitive Analysis

> The biggest problem with current accessibility software is that it is so specialized, and it's hard for people to learn to use it.
>
> - Lorenzo Powell, Job Placement Officer – Center for the Visually Impaired

At this time, no single product provides the integrated, customizable adaptations of which EyeAbility is capable. A table of Crossing Point Technologies' key competitors, their functionality, price, is located at the end of this plan.

In 1998 Microsoft Windows initiated the inclusion of accessibility options into the "Control Panel." In reality, accessing these options can be highly problematic for an individual who has a visual impairment. Furthermore, the solutions provided do not adequately meet the needs of a range of visual impairments. With respect to display, it is possible to change the overall appearance of the display's colors, icon size, cursor blink rate, and cursor size, but only through access of the control panel and through separate functions/menus. The use of these accessibility options can prove awkward and cumbersome for users with visual impairments. The user must know the meaning of each option presented in the accessibility control panel and then determine which options to activate in order to optimize his/her computing experience. EyeAbility shifts this burden from the user to the computer.

In general, the accessibility options are problematic because they are not universal to all of the programs in a computing system, or all components of a program. A user or a third party assisting the user has to manually change the accessibility options through the control panel.

Marketing Strategy

Crossing Point's overall strategy is to reach home users with a low cost strategy to create demand in the business realm. The implementation strategy will be honed over the next six months by existing management and a contract-marketing expert. This person has already been identified. He will begin work May 2005 and is interested in coming on board full time once funds are available. The following is our initial strategy based on preliminary market data and discussions with the advisory board.

> Direct Marketing is an intelligent way to market EyeAbility, as it enables Crossing Point to engage with prospects in every stage of the sales cycle.
>
> - Ann Bachrach, Manager - Babcock & Jenkins

Sales/Promotion

Crossing Point's sales strategy is a two-pronged approach that is focused on a direct mailing campaign to home users followed closely by direct sales to businesses and government.

Consumers **-** Crossing Point reaches consumers using a direct mail/web-based approach that has been designed with guidance from experts in this area. The company will begin with sending out 10,000 trial versions of EyeAbility to a very targeted group of potential customers. These customers are computer users with household incomes greater than $50,000 a year.

They are further segmented by age range, which will initially be focused on customers 40 years of age and older. One-third of the products will be sent to each of three different age groups: 40-54, 55-64, and 65+. Targeting will be honed and adjusted based on response rates from each group.

Crossing Point will also provide incentives to generate customer response and completion of the performance assessment. First, the company will offer respondents the chance to win a sweepstakes. Second, Crossing Point will offer an interesting e-newsletter that positions the company as a source of trusted information, which is important since people buy from people they trust.

Crossing Point is also focused on generating word-of-mouth marketing through relationships with vision centers, advocacy groups, American Association of Retired Persons, and the Department of Labor. Promotional efforts will include on-site demonstrations and point-of-sales marketing. The company will also follow the drug representative model of giving free samples to doctors and clinicians to give to their clients who will most benefit from the technology.

The costs of a direct mail campaign as estimated by Ann Bachrach at Babcock & Jenkins for an initial mailing of 10,000 contributes to the following expenses, shown in Table 6.3. While marginal cost per mailing is initially $2.25, Crossing Point expects this number to fall drastically as the number of mailings increase.

Business/Government – Crossing Point reaches business and government through direct sales channels. Crossing Point’s executive team has begun preliminary talks with interested customers and is serving as the initial sales force. The company is focused on businesses and agencies with a high usage of computers in industries that will most readily understand the value of the product.

These include industries where reducing errors and improving productivity are important, and they include businesses such as call centers, telemarketers, data processing, and computer programming. In addition, the government is currently the largest employer of individuals with vision impairments, making it an attractive market.

Table 6.3. Direct mail costs.

Creative, production, set up costs	$ 45,000.00
Targeted distribution list	$ 7,500.00
Printing and fulfillment (10,000 @ $1.50)	$ 15,000.00
Total	$ 67,500.00

Price

Pricing for EyeAbility is as follows:

Home Version - $30/year – The product is free for 30 days and is licensed for use by one user on one machine.

Pro Version - $70/year – The license is good for one user on any machine so employees could use the product at home for no additional charge. This is important and will drive business sales because when someone has the product on one machine they will want it on others.

Enterprise Version – This will average approximately $15,000 for a 2000 person company. This license will also allow employees to use the product at home.

Implementation

Below is a more in-depth look at several of these operating initiatives.

Product Development – Ongoing research is critical to Crossing Point's continued lead in the marketplace. CSO V. Kathlene Leonard is focused primarily on continuing the research on which Crossing Point's products are based and will continue to conduct cutting edge, product-focused research to provide the foundation for Crossing Point's future product development.

While the current Alpha version of EyeAbility is functional from a scientific standpoint, there are several additional, critical development tasks that must be completed before the product will be ready for launch in the first quarter of 2006. Important product development and maintenance tasks include:

Develop server based diagnostic component- In the Alpha version, the automated diagnostic portion, including the proprietary database of research metrics, is installed on the client machine. For both database security reasons, as well as the value of gathering customer data metrics for ongoing research, the release version of the product will rely on a server based component that Crossing Point will maintain for the diagnostic processing and database lookups. In this way, the database will never be released as part of the product installation.

Create product installer- The Alpha version has no supporting installation and setup functionality, currently relying on a manual process to get the software functioning on a new machine. An installation and packaging routine must be developed and tested to ensure easy installation. Crossing Point will be using the

InstallShield installation environment for this task.

Setup and test server environment-: The server component will be installed on servers in a co-location facility in Atlanta. Once the server component development is complete, rigorous loading and stability testing will be performed in advance of a release to help ensure a smooth launch.

50 user Beta launch- Once the beta development is complete, Crossing Point will launch the offering to an initial group of previously identified customers. The company is already working with individuals at the Center for the Visually Impaired, Georgia Tech, and Internet Security Systems to identify the beta users. Given the "one-time" nature of the assessment component, the Beta will entail a staggered launch to the 50 users over the course of a month. In this way, adjustments can be made during the rollout if issues come to light.

Product Production & Delivery – The EyeAbility software suite is primarily delivered via online download. All installation, training, and support materials are made available digitally through this process, as well as on the website. This is an important feature, as digital delivery enables Crossing Point to present the information to users with vision impairments in an accessible format. A mini-CD based version will exist for distribution through ophthalmologists, low vision centers, and direct mail; however, this is used primarily as a marketing initiative rather then a traditional retail product package. Production costs for these units are discussed in the market section.

Customer and Technical Support – Initial customer and Level I technical the sales and management team, with Level II, will handle support support being handled by the development team. This will enable employees to better understand the market needs and requirements, while minimizing the costs associated with a full time customer support engineer (CSE) initially. In the second quarter of 2007, as product demand increases, the company will hire a full-time CSE dedicated to this function. Continued sales growth will increase the need for individuals in this department.

Management Summary

Crossing Point's executive team brings a wide range of relevant skills and startup experience to the company. Together, they bring experience starting, growing, and successfully exiting investor-funded ventures. They have experience in HCI technology development, and they have expertise in bringing complex new products to market.

Management Team

David Beck – Chief Executive Officer - David Beck brings 13 years of leadership experience in operations, sales, and business development to the Crossing Point Technologies management team. He has opened four new businesses and consulted with entrepreneurs during the start up of five others. Recently, he was the Director of Operations for CRM, Inc. a large restaurant company, where he was responsible for all aspects of operations for nine business units across two states. His duties included full P&L accountability, sales, marketing, and new business acquisitions. David is currently a

Technology Commercialization Analyst for the Technological Innovation: Generating Economic Results (TI:GER) program at Georgia Tech while he is completing his MBA.

Mike Orndorff – Chief Operating Officer **-** Mike Orndorff provides 11 years of software experience in product development and new product introductions to Crossing Point Technologies.

He was previously CEO of MultiMediums, Inc., a web-based application development firm. Mike oversaw and negotiated the sale of his company to MarketingCentral in 2000. He stayed on in the role of VP, Product Marketing, where he managed 2 new product introductions and had responsibility for sales and partnership development. Mike is currently a Business Analyst for the Advanced Technology Development Center (ATDC), a high-tech business incubator. He is pursuing an MBA at Georgia Tech where he also graduated with honors in 1999 with a BS in Electrical Engineering.

V. Kathlene Leonard – Chief Science Officer - Kathlene Leonard is a Ph.D. Candidate in Industrial Systems Engineering (ISyE) at Georgia Tech. Prior to her Ph.D. work; Kathlene received her B.S. in Industrial Engineering from University of Wisconsin. Since January 2001, she has been employed as a Researcher in the Laboratory for HCI and Health Care Informatics (HCI). Her work emphasizes applications and theory development concerning human aspects of personal and networked computing, as well as universal access to information technologies. Mrs. Leonard was recently awarded an NSF IGERT fellowship. This fellowship has led to involvement with the TI:GER program within Georgia Tech's College of Management.

Financial Plan

Financial Projections

Crossing Point has projected EBIT of $8.9 million in 2009 on $13 million revenues. This represents robust growth over the first five years of operations. The company is projected to be profitable by the third quarter of 2007 and reach positive operating cash flows by the fourth quarter of 2007.

First Year

In the first year, Crossing Point is focused on minimizing costs while executing the final development tasks to get the initial product, EyeAbility, to market. By completing product development and proving market acceptance, the company will minimize risk for the Series A investors while minimizing dilution for the founders. Initial financial resources will be provided by principals, during which time additional seed funds will be raised through grants and a "friends and family" round of investment.

Investment Offering

Once the product is complete and initial reference customers are installed, Crossing Point will seek $3,000,000 in Series A financing. The company is looking to receive this funding by the first quarter of 2006. Primary use of funds (for a detailed account see pro-forma financials at the end of this plan) includes marketing and PR campaign as well as salaries and expenses for direct sales.

In exchange for the investment, the investor will receive stock representing a 38% ownership of the company. The remaining 62% of the ownership will be maintained by the founding team and seed-stage investors.

Exit Strategy

As the company grows, steps will be taken to enable investors to participate in an exit to maximize the returns to the investors. The company expects the most likely exit to come in the form of a Merger with an existing accessibility company or Acquisition by a larger computer Original Equipment Manufacturer (OEM).

Merger with Existing Accessibility Company

By year five, the revenues and proven scalability will make the company an attractive partner for a company such as Freedom Scientific, one of the premiere companies in the accessibility industry today. The combined product lines from Crossing Point and Freedom Scientific would create a resulting product suite that would holistically address the vision-impaired market across the entire spectrum of visual acuity. This represents a much larger opportunity then Freedom Scientific is currently able to penetrate today.

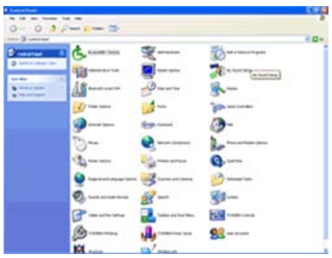
Crossing Point has the added benefit of developing relationships with potential acquirers from the start. The company is working to build partnership agreements from the beginning with existing accessibility software companies, as

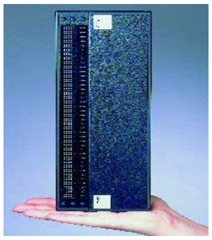
Crossing Point's products serve as a natural sales channel for some of these products.

Acquisition by OEM

The other likely exit scenario would involve a computer manufacturer such as Dell or Gateway. Crossing Point software would serve as an excellent bundled offering to serve as a key differentiator for an OEM when selling to a growing aging population.

6.3 Assistive Technology Competitor Overview

Competitor	Product Overview	Key Limitation(s)	Price
Microsoft Windows Accessibility Options 	Accessed through the control panel, user can turn on several different accessibility features to different levels.	Difficult to access (especially if the user has an impairment) Difficult to know set optimal combination accessibility settings Limited number of changes directed towards visual impairments	Integrated within the Microsoft Windows Platform since 1998
Screen Reader 	Software program that reads to the user elements that appear on an interface via synthesized voice. The program reads left to right, starting at the very top of the screen. When an image is encountered the program reads the associated ALT text.	Solution abandons any remaining vision the user has, using only their auditory ability Efficacy depends on the organization of the interface (e.g., anything not modeled in a left to right organization is not compatible) Cannot be transferred between systems A ‘one-size fits all’ solution	$700 - $1,000

Braille Display 	Similar to the screen readers, but gives the reader the information via tactile cues (Braille characters).	Same limitations as a screen reader, plus the user has to learn Braille, which is not likely if they have residual vision, and are losing vision later in life	$4,000 - $10,000
Screen Magnifiers 	Physical device or software program that enlarges the entire screen image.	A 'one size-fits all' solution, that is not adaptable between users Cumbersome, cannot be transferred between systems For people with visual impairments, magnification is not always the most effective strategy (especially with obstructed visual fields)	$100-$600

6.4 Critical Risks

Critical Risk	Description	Mitigation Strategy
Competition	Microsoft or another large company could decide to enter this market	Crossing Point's core research provides a significant barrier to entry. To effectively bring a duplicate of EyeAbility to market, competitors would have to perform the underlying vision research, which is not what software companies traditionally do. It is more likely that a large company would be interested in acquiring Crossing Point's technology.
Technology	Full product development has not been completed.	The company is doing ongoing testing and research to optimize product design. Alpha results are good and more work is being done. Further, Crossing Point is lowering investor risk by not seeking major funding until the Beta product is functional.
Access to Capital	Minimum requirements for capital must be met to get the venture started. One risk the venture faces is that these requirements are not met up front.	The company will leverage relationships with previous investors, VentureLab, and ATDC to gain access to capital. Secondary funding/bootstrapping plans are in place should initial efforts not bear fruit.
Slow Sales	End users could be slow to adopt the new technology.	Initial development focuses on a product that meets a pressing market need that is easy to use. Additionally, marketing efforts are focused on driving interest to create market pull.

Board of Advisors

Technology Advisor - Dr. Julie Jacko is an Associate Professor of Industrial & Systems Engineering (ISyE) at Georgia Tech and is the author or co-author of over 100 research publications including journal articles, books, book chapters, and conference proceedings. She is also the Director of the Laboratory for Human-Computer Interaction and Health Care Informatics (HCI) in ISyE. Dr. Jacko's research activities focus on human-computer interaction, human aspects of computing, and universal access to electronic information technologies. Her externally funded research has been supported by the Intel Corporation, Microsoft Corporation, the National Science Foundation, and NASA. Dr. Jacko received a National Science Foundation CAREER Award for her research titled, "Universal Access to the Graphical User Interface: Design For The Partially Sighted," and the National Science Foundation's Presidential Early Career Award for Scientists and Engineers (PECASE), which is the highest honor bestowed on young scientists and engineers by the federal government.

Marketing Advisor – Ann Bachrach is an advertising professional with over 12 years of marketing experience. Formerly a Partner in a boutique advertising agency in Atlanta, Georgia, she served clients such as The Coca-Cola Company, BellSouth, MCI, CNN and Delta Air Lines. She has recently moved to Babcock & Jenkins, a direct and relationship marketing agency, where she focuses exclusively on high-tech clients that include Polycom, Symantec and Nortel.

Vision Industry Advisor – Subie Green is the Executive Director of the Center for the Visually Impaired, Georgia's largest comprehensive, fully accredited, private facility providing rehabilitation services for individuals of all ages who are blind or visually impaired. During her tenure as Executive Director, the Center has grown to serve as a model of innovative services for people who have a wide range of vision impairments from low vision to total blindness.

Commercialization Advisor - Steve Derezinski is the Director of VentureLab at Georgia Tech. As director, Mr. Derezinski is responsible for the day-to-day operations of VentureLab and works with faculty to help commercialize their research at Georgia Tech. Prior to this position; he was founder and CEO of SmallBizPlanet.com. He is advising Crossing Point Technologies on getting the technology from the laboratory to the marketplace.

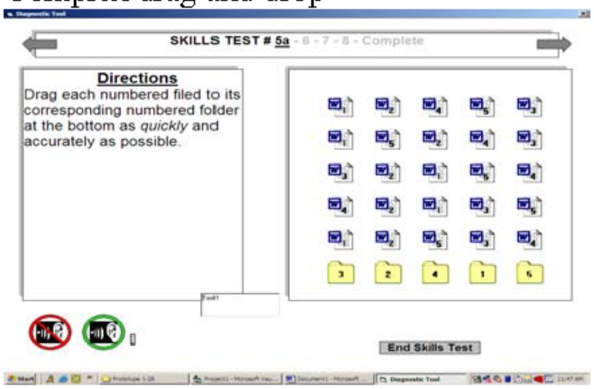
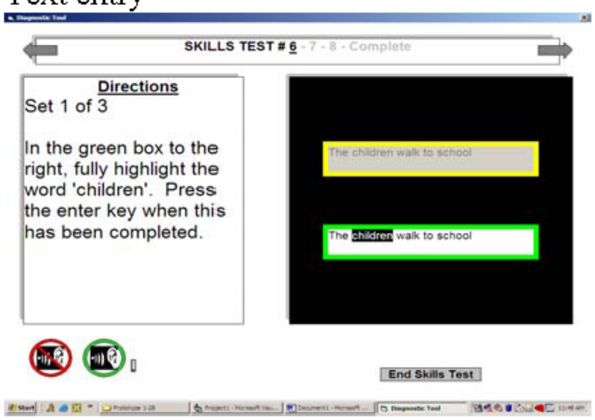
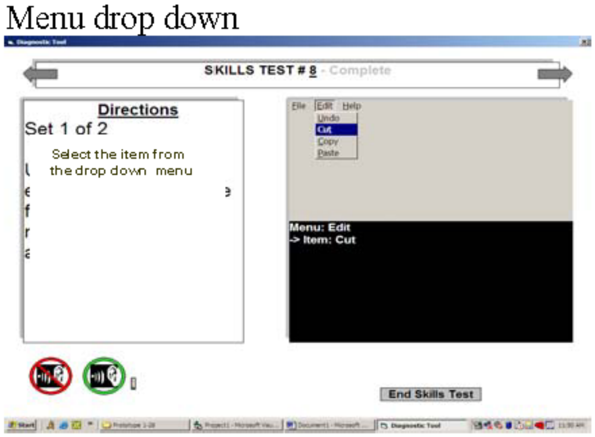
Strategic and Operations Advisor - Tony Antoniades is the General Manager of the Advanced Technology Development Center (ATDC), a nationally recognized science and technology incubator that helps Georgia entrepreneurs launch and build successful companies. In his role as General Manager, Tony is responsible for the overall ATDC strategy, products and services.

Legal Representation – Charles Vaughn is a partner of Nelson Mullins in Atlanta where his practice focuses on corporate finance and related matters. Mr. Vaughn represents public and emerging growth companies in private and public offerings of securities, from early stage offerings made to friends and family or angel investors, to one or more venture capital rounds, to initial public offerings and secondary offerings after companies are public. Mr. Vaughn also represents companies in mergers and acquisitions.

Business Advisory Services – HLB Gross Collins provides accounting and business consulting services for Crossing Point.

6.5 EyeAbility Alpha Screen Shots

A portion of the diagnostic tests and metrics, are shown below, as they appear in the current version. The application automatically steps a user through tasks, giving both visual and verbalized auditory instructions.

Complex drag and drop 	Clicks on file, clicks off file, target re-entries, drop positioning errors, traveling drop errors, task axis crossings, vertical movement direction change, orthogonal direction changes, movement variability, errors, movement offset, time components
Text entry 	Time per sentence, words per minute, wrong order, wrong words, missing words, extra words, changes made, highlight commissioned, wrong characters, time per word, other time components
Menu drop down 	Time per selection, time to correct menu, wrong item selected, wrong menu selected, number of incorrect selections, missed opportunities, number of items highlighted, menu position, other time and error components

VITA

V. Kathlene (Emery) Leonard was born in 1977 in Cape Girardeau, Missouri and grew up in Janesville, Wisconsin. Katie attended the University of Wisconsin-Madison, graduating in 2000 with a B.S. in Industrial Engineering. At the University of Wisconsin-Madison, she began graduate studies in the Industrial Engineering and Human-Computer Interaction under the direction of Dr. Julie Jacko. In January 2001, she transferred to the School of Industrial and Systems Engineering at the Georgia Institute of Technology, in Atlanta, Georgia. There, she continued her work with Dr. Jacko as Research Assistant in the Laboratory for Human-Computer Interaction and Health Care Informatics (HCI).

Katie’s work has emphasized applications and theory development concerning human aspects of personal and networked computing, as well as universal access to information technologies. Katie's specific research interests involve theoretical development and empirical validation of constructs, frameworks, models and metrics related to the evaluation and design of accessible information technologies.

While at Georgia Tech, Katie was recognized as National Science Foundation Integrative Graduate Education and Research Traineeship (IGERT) Fellow, Georgia Tech Presidential Fellow, and was the recipient of an International P.E.O. Scholar Award in 2004-2005. After graduations Katie will complete a post-doctoral appointment in Biomedical Engineering at Georgia Tech, ultimately pursuing a career in academia.