Behaviour & Information Technology, Vol. 25, No. 4, July–August 2006, 313–332

An investigation of handheld device use by older adults with age-related macular degeneration

V. K. LEONARD†, J. A. JACKO*‡ and J. J. PIZZIMENTI§

†Alucid Solution, Inc., Atlanta, Georgia, USA
‡Wallace H. Coulter Department of Biomedical Engineering, Georgia Institute of Technology and Emory University School of Medicine, 313 Ferst Drive, Atlanta, GA 30332-0535, USA
§College of Optometry, Nova Southeastern University, Fort Lauderdale, Florida, USA

This study investigates factors affecting handheld human–computer interaction (HCI) for older adults with Age-related Macular Degeneration (AMD). This is largely an uncharted territory, as empirical investigations of HCI concerning users with visual dysfunction and/or older adults have focused primarily on desktop computers. For this study, participants with AMD and visually healthy controls used a handheld computer to search, select and manipulate familiar playing card icons under varied icon set sizes, inter-icon spacing and auditory feedback conditions. While all participants demonstrated a high rate of task completion, linear regression revealed several relationships between task efficiency and the interface, user characteristics and ocular factors. Two ocular measures, severity of AMD and contrast sensitivity, were found to be highly predictive of efficiency. The outcomes of this work reveal that users with visual impairments can effectively interact with graphical user interfaces on small displays in the presence of low-cost, easily implemented design interventions. Furthermore, results demonstrate that the detrimental influence of AMD and contrast sensitivity on handheld technology interaction can be offset by such interventions. This study presents a rich data set and is intended to inspire future work characterizing and modeling the interactions of individuals with visual impairments with non-traditional information technology platforms and contexts.

Keywords: Older adults; Visual impairment; Macular degeneration; Icons; Drag and drop; Spacing; Auditory feedback; Mobile computing; Handheld computers

1. Introduction

Over the past 10 years, a growing body of research has focused on understanding and improving access to Information Technology (IT) for individuals who experience some level of visual dysfunction. This is largely motivated by an expanding population of older adults, as it is estimated that 1 in 3 baby boomers will experience a vision-reducing eye disease by the age of 65. By the year 2030, the population of Americans 65 and older will number 70 million (Quillen 1999), generating an urgency for advancements in accessible technology for this population.

Previous work has demonstrated that interactions are strongly influenced by the nature and amount of residual vision a user possesses in combination with the computer interface characteristics (summarized by Jacko *et al.* 2005). This underlying concept has spawned several theories of IT interaction for individuals with visual impairments.

- IT solutions for individuals who are blind are typically inappropriate for individuals maintaining useful residual vision possessed by the user.
- The efficacy of design interventions depends on the nature and amount of a user's residual vision.

---

*Corresponding author. Email: Jacko@bme.gatech.edu

Behaviour & Information Technology
ISSN 0144-929X print/ISSN 1362-3001 online © 2006 Taylor & Francis
http://www.tandf.co.uk/journals
DOI: 10.1080/01449290600636512

- Increasing text size and image size can be more problematic than assistive, especially considering the nature of the visual impairment.
- The emphasis of direct manipulation tasks on visual interaction paradigms places users with visual impairments at a quantifiable disadvantage when attempting to use graphical user interfaces (GUIs).

The present study aims to further expand our understanding of human–computer interactions (HCIs) for older adults with visual impairments, through an appraisal of influential factors of direct manipulation on a handheld computer. The outcomes of this work serve to establish a much needed baseline assessment of this population's interaction with handheld devices for future researchers to consider in light of the complex contextual and task-related aspects of this type of interaction.

1.1 *Icon manipulation and visual impairments*

The work presented in this manuscript is a component of a much larger research agenda aimed at the empirical characterization of the manipulations critical to GUI interaction for a variety of users who represent a range of visual profiles (visual functioning capability and diagnoses). To this end, Jacko and colleagues have completed several empirically based studies demonstrating how characteristics of users' functional vision and ocular diagnoses can influence performance on a variety of direct manipulation tasks for effective GUI interaction (e.g. Jacko 1999, Jacko *et al.* 2000a,b, 2002a,b, 2005, in press).

These studies have addressed the relative performance of a cohort of users with visual impairments due to ocular disease and a cohort of age-matched controls without ocular dysfunction on several desktop computer tasks. Assessments of the interactions are achieved via traditional time and accuracy measures of performance, but also physiological methods such as electroencephalogram (Jacko *et al.* 2000a) and eye tracking (Jacko *et al.* 2000b). HCI issues considered relative to visual profile include:

1) the visually rigorous task of icon search and selection in the presence of distracters (Jacko 1999, Jacko *et al.* 1999, 2000c, 2001, 2002a, Scott *et al.* 2002a,b);
2) cursor movement (Jacko *et al.* 2000c);
3) the direct manipulation of drag and drop in the absence of distracters (Jacko *et al.* 2002b, 2003, 2004a, 2005); and
4) the identification and selection of targets in a drop-down menu with distracters (Edwards *et al.* 2004, 2005; Jacko *et al.* in press).

The emergent theme, which links these research studies, is the importance of understanding the specific details of a user's impairment in terms of their functional ability to achieve reasonable levels of performance in HCI tasks. In addition, these studies measured the impact on performance of enhancements to the visual display (e.g. increasing contrast, size, and altering color), and also the influence of augmenting the visual display with supplemental multimodal feedback (i.e. auditory and/or haptic cues).

Based on the work completed to date by the authors (for a complete review, please see Jacko and Leonard (in press)), the current study focused on a complex drag and drop task integrating auditory feedback on a small, mobile display. This work both adds incrementally to this subject and contributes in novel ways with its involvement of a handheld interactive environment. Synopses of select studies from this body of work most relevant to the present study of handheld HCI are subsequently provided.

1.1.1 Iconic visual search. In the examination of iconic visual search (such as file, print, save, etc.) in the presence of distracters, visual acuity, contrast sensitivity and color perception were found to be significant predictors of performance on this search task for users with Age-related Macular Degeneration (AMD) (contrast sensitivity was the most sensitive indicator) (Jacko *et al.* 2001). That is, aspects of visual function affected the performance of various task components differently. In addition, icon size, set size, and background color significantly influenced interaction as a function of ocular diagnosis.

1.1.2 Icon manipulation. A later study of GUI iconic manipulation again considered a population with AMD (Jacko *et al.* 2002b, 2005). Working on a desktop, participants with AMD were tasked with selecting, dragging, and dropping a single Microsoft Word Windows® file into a single Microsoft Office Windows® folder icon. This study measured the efficacy of supplemental multimodal feedback (haptic, auditory and visual) for participants possessing different visual acuities. Unlike previous studies, this experimental task did not have a substantial visual search component, but instead focused on the physical manipulation of the icons on the display. Results demonstrated significant differences in performance between groups of people with different visual acuities on task time, feedback exposure times, and the frequency of errors. Performance improvements were realized for both visually healthy and AMD participants when provided with non-visual and multimodal feedback. Effects were greater in magnitude for participants with the most severe vision loss and AMD.

1.1.3 Additional studies of HCI and visual impairment. While the foundational research described above is one of the largest empirical endeavors to understand the HCI of this population, 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) finds that the majority of knowledge (with rare exceptions) resides in (1) the magnification of screen elements (Fraser and Gutwin 2000); (2) the accessibility of text (Craven 2003); and (3) assisted navigation through websites and other web-based information (Arditi 2003).

In their seminal work, Kline and Glinert (1995) presented UnWindows V1, a set of interface tools to support selective magnification of a 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 reactions 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 reported finding this annoying). Also, 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 (2000) discuss the impasses imposed by the mouse pointer to direct manipulation for individuals with visual impairments. Visual impairments create barriers that diminish the ability to distinguish fine details of iconic screen targets, as well as the ability to track the highly dynamic nature of the pointer used to manipulate these icons. The authors attribute the difficulty in manipulating objects with the pointer to reduced visual acuity and constrained visual field on the basis of four dimensions: (1) *mode*—the sensory channel through which assistance is provided to the user; (2) *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; (3) *dependence*—how the pointing solution, interface, and the onscreen pointer are interconnected; interface dependent or independent; and (4) *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 of most direct manipulations with GUIs employing the Windows-Icon-Menu-Pointer (i.e. *WIMP*) interaction paradigm.

Arditi (2004) addresses the reading difficulties of individuals with visual impairment. 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 acuity.

Findings from these and other studies have established a baseline understanding of HCI for users with visual impairments. They differ from the contributions of the present work, in that most publications on this topic are largely qualitative or design driven, rather than empirical in nature. Still, neither research paradigm has fully considered HCI and GUI design outside the context of desktop computing, in light of the explosive potential for ubiquitous computing. Particularly, there is an absence of investigations that consider the enabling and disabling facets of mobile devices for individuals with visual impairments.

1.2 *Mobile computing and visual impairment*

Researchers have only recently started to ask questions concerning the use of mobile, wireless technologies by mainstream users, let alone by users with limited abilities such as visual impairment. Mobile computing introduces new challenges by providing powerful computing behind suboptimal interfaces: small visual displays, poor audio facilities and limited input techniques. Interactions with mobile computers are also susceptible to the effects of context: varying tasks, environments and users. Users with visual impairments who wish to use mobile computing technologies, such as cell phones and handheld computers, are likely to encounter these contextual challenges in addition to barriers of interaction imposed by their functional vision, or disability-induced impairments (Sears *et al.* 2003, Barnard *et al.* 2005, in-press). Tasks such as way-finding, memory recall and communication can be enhanced for this population with mobile devices, but only if the effects of context and visual ability are adequately accounted for.

As a starting point, this study applies the research methodologies proven successful in the assessment of the impact of impairment on GUI interaction in a desktop environment to direct manipulations using a mobile device, specifically a handheld computer. While the effects of context and task are especially important for consideration in mobile HCI research, work that establishes a baseline performance capacity is a critical first step. In particular, this baseline data is necessary to normalize the data of future studies in which the task, environment, and user profile are more pragmatically dynamic. It can serve to prioritize development of accessible applications for this population. This study aims to reveal, for older adults with AMD, the personal, ocular and interface factors that influence different components of the handheld HCI.

1.3 *Age-related Macular Degeneration*

This study considers the handheld HCI of users diagnosed with AMD. Aging is connected with natural declines in a person's sensory abilities. As such, age is often accompanied by changes to the eye, including the retina and visual nervous system, impacting 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.

AMD is the leading cause of vision loss in adults 55 years and older, and affects more than 10 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). AMD is a disease that affects the center of the retina, or macula, roughly 3 mm in diameter (see figure 1). This portion of the retina is primarily responsible for central, fine detail, and color vision. Accordingly, individuals diagnosed with AMD often experience measurable distortion or deficits to their central visual field, while the vision in their periphery remains intact. This intact vision is referred to as residual vision (Kaufman and Alm 2003). AMD entails a progressive deterioration to the central, high-resolution vision, which over time reduces the 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). Figure 1 provides a simulation of the extent of the impairment imposed on patients at various stages of AMD.

Figure 1. 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).

Accurate diagnosis of AMD is achieved via ophthalmic examination of the posterior of the eye. Visible features on the retina facilitate the diagnosis and classification of AMD. Experts scan the retina for the presence of drusen—discrete yellowish-white spots on the image. In addition they examine the state of the retinal pigment epithelium (RPE), a single layer of cells between the retina and the underlying blood vessels.

Several classification systems have been used to grade the severity of AMD, once it has been diagnosed, all derived from an ophthalmic exam. This study employs a method introduced in 1989 (Bressler *et al.* 1988), which involves grading the severity level on a scale from 0 (no disease) to 4 (most severe) based on the amount of drusen, their distribution on the macula and the observed condition of the RPE. Grade 4, the most severe or final stage, is assigned to those cases in which the RPE is deteriorating or leaking (which can cause a blind spot in the visual field of that eye). The detailed protocol for scoring is referred to in the ophthalmology community as the Chesapeake Bay scoring system. This system was found to be superior in diagnosis of severity to other methods, such as those based on visual acuity (ability to resolve fine detail) (Bird *et al.* 1995).

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 1500 μm of the foveal center (center of the retina), or at least 10 small drusen between 1500 μm and 3000 μm of the foveal center;
- *Grade 2*: many small drusen $\sim$20 or more, within 1500 μm of the foveal center;
- *Grade 3*: eyes with large confluent drusen or eyes with focal hyperpigmentation of the RPE; and
- *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 Nova Southeastern University (NSU) team of optometrists and technicians rated the type of AMD present in each eye as 'Wet' or 'Dry' according to any visible leakage discerned on the RPE.

AMD seldom causes complete vision loss, leading these individuals to adaptively rely on their useful residual vision.

As their vision diminishes, people with AMD learn to integrate non-visual cues with the residual vision. The HCI needs of this user group are significant because those who acquire AMD are likely to experience increases in severity level and associated declines in visual function over time. There is no known cure for AMD. Those with the condition manage the impact of this disease on activities of daily living by developing strategic coping skills; altering behaviors and making use of assistive devices to maintain independence.

1.4 *Study objectives*

The objective of this study is to identify indicators which predict successful iconic search and manipulation using a handheld computer for older adult users with AMD. This investigation considers demographic characteristics of the user, clinically acquired ocular measures and features of the interface and task. This paper reports on three time-based measures (trial time, visual search time, and movement time) and one distance-based measure (icon drag distance). Finally, participants' reactions to the handheld device and the task are summarized through their responses to an exit survey.

Three design interventions were investigated as well as the statistical interactions between each intervention and the severity level of AMD for each were addressed. The interventions included two factors related to screen real estate: the set size (number of icons on the screen) and inter-icon spacing. The third factor is the presence (or absence) of auditory feedback. These factors were considered independently in both mainstream HCI investigations on icons and drag and drop (e.g. Brewster 1998, Harper *et al.* 2001, Vitense *et al.* 2003, Everett and Byrne 2004), but also in investigations targeting populations with visual impairments (e.g. Jacko *et al.* 2000c, 2002a, 2005, in press).

2. Methodology

Thirteen volunteers from the Nova Southeastern University (NSU) College of Optometry patient pool and associates of NSU staff participated in the study. Ten participants were diagnosed with some level of AMD, while the remaining three were visually healthy, age-matched controls. Criteria for inclusion in the study were computer experience (frequency of use and/or application familiarity), age (over 50 years), and ocular diagnosis.

The computer experience survey employed in the study was derived from Emery and colleagues (Emery *et al.* 2003) and allots a score based on the sum of frequency of use and the number of computer-based applications the participants are familiar with. The minimum score for inclusion in this study was 3, mean score (standard error) of the participants was 9.1 (0.192) and the median score was 10. This method allows for the inclusion of individuals who may have used a computer regularly at one time, but have had to abandon use due to their visual impairment.

Controls were included based on the absence of ocular pathologies, while the AMD participants were screened to confirm the diagnosis of AMD and absence of other ocular pathologies. As incentives, participants were provided with comprehensive ophthalmic exams and given $50 US. When necessary, participants were provided with temporary frames outfitted with corrective lenses to enable use of their best-corrected vision for the handheld experimental task.

Participants' self-perceived assessment of health was measured using the SF-12, which generates scores for both mental and physical health (Ware *et al.* 1995). The manual dexterity of the participants' dominant hand (used to control the stylus for the handheld input) was measured with the Purdue Pegboard test of manual dexterity (Tiffin and Asher 1948). Participants' perception on the impact of their visual function on the quality of life in daily tasks was captured with the National Eye Institute Visual Functioning Questionnaire, 25-question version (Mangione *et al.* 2001), reported in the analyses of the exit surveys. While none of the participants had previous experience with handheld computers, nine owned cell phones for at least two years.

Participants interacted with a Dell Axim™ X30 Pocket PC. The handheld display was a touch-sensitive LCD, measuring 3.5 inches diagonally, with the resolution set to 240 × 320 at 16-bit color. The device was secured to an inclined platform during the task to accommodate the collection of eye movement data (reported in a subsequent paper), shown in figure 2. Participants were seated a comfortable viewing distance from the handheld device and

Figure 2. Handheld experimental configuration including screen shot of the task (not actual size).

allowed to adjust the seating for their own comfort. Upon completion of the tasks, participants were verbally administered an exit questionnaire requesting their opinion on a variety of aspects of the experiment, technology, and their perceptions.

The experimental task was designed to assess a range of iconic manipulations and the associated difficulties imposed on this population's interactions with handheld computers. The task required visual search for a target icon among distracters, selection of the icon with the stylus, and finally the drag and drop of the icon to a new target location. In contrast to the Microsoft Word® icons used in previous studies (Vitense *et al.* 2002, Jacko *et al.* 2005), this study used icons of playing cards as the target icons (shown in figure 2).

While participants were screened for computer experience, the majority of their experience had been derived from Internet use, email and games. The use of the file and folder icons may have caused individuals with greater amounts of computer experience or experience with certain applications to interact at higher rates of efficiency due to their familiarity and comfort with the images. The playing cards were more likely to be highly familiar images for a greater number of participants, because a large number of older adults play card games on a regular basis (it has been shown to mitigate effects of aging and dementia; Coyle 2003).

The design elements of the card icons embody the criteria for simple icons of good quality. That is, icons discriminated by as few features as possible, using simple shapes and colors (Everett and Byrne 2004). Decreasing icon quality has been shown to cause inefficient, longer visual search strategies for the visually healthy population, particularly as the number of distracters competing with the target icon increase. The use of playing cards provides some control over the factors of icon quality and familiarity while isolating factors affecting visual search and icon manipulation.

A custom software application was written for this experiment using Visual C. The playing card icons used in the study were numbered from 2 to 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 (i.e. ♥ ♦ ♣ ♠), in their traditional red and black colors. The icons were consistent in size with the standard Microsoft Windows Mobile 2003® icon size, 32 × 32 pixels and appeared to the users as 7 × 7 mm on this display.

Participants were verbally instructed to locate a target card amongst a grid of several distracter card icons of different numbers and suits, select the target using the stylus, then 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. Participants were directed to work as quickly and accurately as possible. Before commencing the trials, participants were trained on the task, informed of the upcoming changes to the interface and introduced to the auditory feedback for the task (volume levels adjusted for adequate detection).

Three independent variables were controlled during this task: Set Size (SS), Inter-Icon Spacing (ISp), and Auditory Feedback (AF).

- *Set Size (SS)* was defined in this study as the number of icons in the playing card grid, that is, the target icon plus the number of distracter icons. For the present study, the SS levels were considered purely on the basis of the screen real estate available on the Pocket PC. Three levels were considered: 4, 8, and 12. The card icons were always distributed four per column, for one, two and three columns in the respective conditions.
- *Inter-Icon Spacing (ISp)* was the distance or white space between the card icons and drop piles measured relative to icon size. While ISp has not been considered before in assessments of interactions for users with visual impairments, it has been shown to be influential in visual search and icon manipulation for a visually healthy population (Hornof 2001); objects near the target were observed to affect the search and selection of the object. ISp had three levels in this study, also based on the limits of screen real estate. The levels include ¼ icon width (1.75 mm), ½ icon width (3.5 mm) and 1 icon width (7.0 mm).
- *Auditory Feedback (AF)* was an auditory cue indicating that a card icon was in position for a successful drop into the pile. If the stylus was lifted at the time the sound occurred, the card would effectively drop into the pile, completing a single trial. Levels of AF were varied between present and absent.

Previous work with non-visual auditory cues and drag and drop employed an auditory icon, a 'sucking' sound to signify accurate placement for releasing the file into a folder icon (Jacko *et al.* 2005). The present study employed the same auditory icon as employed by Jacko and colleagues. However, the previous study applied the auditory feedback in a task where the display comprised a single file and a single folder. The present study introduces the effects of distracter icons and target destinations for the drop.

The factorial design generated for the present study (3 × 3 × 2) resulted in 18 total interface conditions with nine repetitions. Twelve participants completed all 162 trials, and one completed 93 trials. The order of participant exposure to the 18 interface conditions was divided into two sets: AF present and AF absent. The conditions within AF present and AF absent were completely randomized and the order of exposure to the AF sets was random across all of the participants.

The arrangement of the card icons, drop piles and the collection of distracter card icons were randomly assigned for each trial across the participants. The target card for each trial was consistent between the participants for simplification of the experimental protocol. While they searched for the same target cards at trial 1, 2, and so on, the conditions under which they sought that icon differed to mitigate any specific impact of card number or suit.

The dependent variables profiled the overall efficiency and effectiveness of the interaction and accounted for several subcomponents of the task. This paper reports four continuous measures of performance: three time-based (measured in ms), and a fourth measure, which was the distance (pixels) traveled with an icon, prior to its final release into a drop pile.

- *Trial Time (TT):* a measure of the total time from first exposure of the task screen until a card icon (not necessarily correct) is dropped into one of the card piles (not necessarily correct).
- *Visual Search Time (VST):* a measure of the time between when the task screen first appears, until the stylus touches the active area of the icon which is ultimately dropped into a pile.
- *Movement Time (MT):* based on the icon that is ultimately dropped into a pile, this is the time between when the user first selects the card using the stylus and lifts it from the screen to when it is successfully dropped into a pile.
- *Drag Distance (DD):* the number of pixels over which the stylus was used to drag the card icon before its successful drop into a pile. A greater DD can indicate a lack of efficiency in the card movement to the pile.

3. Results

Overall, participants demonstrated a high rate of accuracy during task completion (97% of trials resulted in the correct card being dropped into the correct pile). Linear regressions were applied to the time and accuracy data to ascertain the most influential factors on handheld HCI for this population. The report of the linear regression analyses is followed by a summary of participants' subjective responses collected through the exit surveys.

3.1 *Performance analyses*

The utility of regression in explaining human–computer interactions in related contexts of use was demonstrated by Edwards and colleagues (2005) in an assessment of sources of performance variability for users with diabetic retinopathy performing a drop-down menu task. The sources of variability considered in the present study are summarized in table 1, classified according to interface, participant and ocular characteristics (ocular health and function). In addition to those variables listed in table 1, statistical interactions between the AMD severity score and the independent factors were introduced into the models. It should be noted that participants' perceptions of visual functioning, measured with the National Eye Institute VFQ-25, was not included as a predictor in these models, due to the high correlation of this outcome measure with the various clinically acquired ocular measures. This high correlation violates the assumption of no colinearity between predictor variables in analyses of regression. The collection of predictor variables entered into each regression was consistent, enabling comparisons of the relative effects of the predictors within and between the models.

While analyses using group comparisons have more commonly been employed in previous research involving HCI and visual impairments (e.g. Jacko *et al.* 2002a, 2004b, 2005), the regression approach was applied to this data set for several reasons. 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 grading AMD severity (Bird *et al.* 1995). Jacko and colleagues (Jacko *et al.* 1999, Jacko 2000) identified several visual factors that influence performance during visual search for an icon, including visual acuity, contrast sensitivity, and visual field. In consideration of the great number of measures taken to profile the various phases of the interaction, both implicit and explicit, regression provides a more suitable means by which to compare the relative influence of the various independent variables in relation to each other. More specifically, regression enables the exploration of the impact of the predictors on the components of the interaction relative to each other (Field 2000), and is an especially useful statistical tool in investigations involving novel constructs (i.e. handheld platform).

The utility of regression in explaining HCIs and visual ability was demonstrated by Edwards and colleagues (2004, 2005). In this work, regression was used to assess performance variability for users with diabetic retinopathy using a drop-down menu under various interface conditions. Further, Scott *et al.* applied regression modeling to their examination of the factors affecting icon recognition and selection (2002a,b).

While these analyses may at first seem overly involved for a seemingly simple experimental design, they are necessary to account for the inherent heterogeneity of this participant

Table 1. Predictor variables considered.

Predictor	Description	Observed levels
Interface-related characteristics		
Set Size (SS)	The number of card icons presented for each trial	1 = 4 card icons 2 = 8 card icons 3 = 12 card 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	The column where the target card icon is located for each trial	1 = leftmost 2 = middle 3 = rightmost
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	Sequential position of the trial within a participant's overall experimental session	Range: 0–161
Age	Age of the participant	53–82 years Mean = 68.69 years Median = 70 years
General participant-related characteristics		
Physical Health (PCS)	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
Mental Health (MCS)	Self-reported mental health at the time of the experiment, from 1 (worst) to 100 (best)	Range: 26.39–60.79 Mean = 46.74 Mean = 48.61
Manual Dexterity	The average number of pins inserted into small holes in a board over three, 30 s trials, from 0 (worst) to 30 (best)	Range: 4.67–16.33 Mean = 11.49 Median = 12.33
Ocular-related characteristics		
LogMar Near Visual Acuity[†] (NVA)	Ability to focus on fine details at a distance of 40 cm, translated from Snellen acuity (e.g. 20/20) from 0.1 (best) to 1 (worst)	Range: 0.19–1.00 Mean = 0.71 Median = 0.80
Contrast Sensitivity[†]	Measure image visibility is before it is indistinguishable from a uniform field, from 0 (low) to 60 (high)	Range: 26.00–40.50 Mean = 33.50 Median = 34.50
AMD Severity Score[†]	A diagnosis of severity of disease from no disease (0) to severe (4)	Range: 0–4.00 Mean = 1.17 Median = 1.00

[†] For NVA CS and AMD Score, weighted average of the best and worst eye $(0.75 * \text{best} + 0.25 * \text{worst})$.

population. Cognitive, physical, and perceptual faculties diminish at rates that can vary dramatically between individuals. While selectivity in participant recruitment and careful experimental control can moderate some of these possible outcomes, older adults vary across too many factors for it to be reasonable to wholly account for the lack of homogeneity with simple statistical procedures. It is thus reasonable to look to more complex statistical analyses, such as the regression analyses applied here, which can account for these anticipated co-variates, and even provide additional insight into how they influence the participants' performance.

For each model, stepwise regression was applied to analyze the contributions of the identified predictors to the overall variance of each dependent measure and to identify a linear model that best fit the data. Stepwise regression methods have, in other statistical analyses and discussion, been identified as optimal for exploratory studies, in which little or no previous research exists in the area (Fahrmeir and Tutz 1994, Field 2000). In order to meet the assumptions of regression analysis, transformations were applied to each measure of efficiency and outlying cases were removed to strengthen each model. 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 and 58% of the variability (see table 2).

Table 3 provides a detailed synopsis of each model, including the significant variables, coefficients and standardized coefficients. While the coefficients and constants are beneficial to constructing predictive equations for each variable, the practical interpretation of coefficients is less straightforward, due to the discrepancy in the scales used to measure each predictor variable. The standardized coefficient (B-std) proves extremely beneficial in interpreting the models. It provides the means by which to quantitatively compare the relative impact of each predictor on the efficiency measures within and between models.

Although the values in table 3 are rich with information that is useful in predictive modeling of task efficiency, it is difficult to glean the most salient trends emerging. To this end, figures 3(a–d) provide an illustrative summary depiction. For each model, a bar graph plots B-std for the variables included. By plotting the standardized B-std, relative comparisons can be made in terms of 'how much more' a predictor influences a given efficiency measure, and also enables drawing comparisons between models.

The following should be considered with respect to figures 3(a–d):

- *bars extending to the left of the origin:* an increase in the value of that predictor in the model imposes a decrease in the value of the efficiency measure;
- *bars extending to the right of the origin:* an increase in the value of that predictor imposes an increase in the value of the dependent efficiency measure;
- *increased* $1/\sqrt{TT}$ *and* $1/\sqrt{VST}$*:* faster times, improved efficiency;
- *increased ln MT:* longer icon movement times, degraded efficiency; and
- *increased ln DD* equates to longer distances traveled with the icon for declines in efficiency.

Table 2. Model summary, all models significant at $p < 0.001$.

	$1/\sqrt{TT}$	$1/\sqrt{VST}$	ln MT	ln DD
N	2011	2011	1990	2004
R^2	0.580	0.518	0.487	0.473
R^2-adj	0.578	0.515	0.485	0.470

For example, in figure 3(a), the *AMD Score* bar extends far to the left. This means that as AMD severity score increases (the severity worsens) the anticipated value of $1/\sqrt{TT}$ decreases substantially more than it would in the influence of any other predictor variable. This suggests that AMD interferes with the timely completion of GUI interactions on a handheld device far beyond the interaction of older individuals who do not possess any ocular pathology.

3.2 *Exit survey, subjective participant responses*

After completion of the tasks, participants were asked a series of questions regarding their experiences. Included were questions concerning perceptions of performance and workload during the task, comfort with the equipment, and opinions of the various interface manipulations. The exploratory nature of this study mandated queries of participant reactions to the relatively novel interaction paradigms and hardware configurations.

Participants were, overall, positive about their experiences. With respect to their use of the handheld computer, participants were asked to rate their comfort as very comfortable, comfortable, neither comfortable nor uncomfortable, uncomfortable, or very uncomfortable. The majority rated themselves as very comfortable ($n = 10$), two participants rated themselves as comfortable, and one individual gave a rating of neither comfortable nor uncomfortable. None of the participants elected to report their experience as uncomfortable or very uncomfortable. This dispels, to some degree, the myth that older adults and older adults with visual impairments are uncomfortable with novel interaction paradigms and hardware.

Participants were also asked to verbalize what they liked and disliked most about the task on the handheld device. Comments were varied; in general participants' comments were positive with respect to the task and technology. Several participants responded with *nothing*, to both the question of likes or dislikes, or they referenced an aspect of the experiment such as having to wear a head-mounted eye-tracking unit. Table 4 provides a transcription of individual participant's verbalized likes and dislikes, which were task- and interface-specific.

The positive effect of using the playing card icons is observed in the 'likes' category of table 4. In this category,

Table 3. Summary of predictors and coefficients included.

Variable		$1/\sqrt{\mathrm{TT}}$	$1/\sqrt{\mathrm{VST}}$	ln MT	ln DD
Interface-related characteristics					
Set Size (SS)	B	−0.0016	−0.0029	*****	*****
	SE	0.00010	0.00020		
	B-std	−0.32	−0.34		
Inter-Icon Spacing (ISp)	B	*****	*****	*****	0.244
	SE				0.011
	B-std				0.374
Column	B	*****	−0.00081	*****	*****
	SE		0.0000095		
	B-std		−0.13		
Row	B	*****	*****	0.12	0.45
	SE			0.012	0.013
	B-std			0.18	0.53
Drop Location	B	*****	*****	−0.057	−0.027
	SE			0.007	0.0079
	B-std			−0.140	−0.055
Trial #	B	0.000017	*****	−0.002	*****
	SE	0.0000013		.00016	
	B-std	0.134		−0.17	
General participant-related characteristics					
Age	B	0.000093	0.00012	−0.008	*****
	SE	0.0000086	0.000015	0.001	
	B-std	0.213	0.17	−0.18	
Mental Health (MCS)	B	0.00010	0.00019	−0.010	*****
	SE	0.0000093	0.000017	0.001	
	B-std	0.25	0.27	−0.24	
Physical Health (PCS)	B	−0.000037	−0.00012	*****	*****
	SE	0.0000081	0.000015		
	B-std	−0.085	−0.16		
Dexterity	B	0.00036	0.00056	*****	−0.031
	SE	0.000035	0.000064		0.0035
	B-std	0.26	0.24		−0.17
Ocular-related characteristics					
Near Visual Acuity (NVA)	B	−0.0033	*****	0.89	−0.18
	SE	0.00033		0.040	0.048
	B-std	−0.22		0.53	−0.088
Contrast Sensitivity (CS)	B	0.00038	0.00031	−0.081	0.018
	SE	0.000027	0.000039	0.003	0.0037
	B-std	0.37	0.18	−0.72	0.13
AMD Score	B	−0.0028	−0.0049	0.17	0.078
	SE	0.00016	0.00028	0.017	0.011
	B-std	−0.72	−0.75	0.38	0.15
Interaction terms					
AMD*SetSize (AMD*SS)	B	0.00027	0.00057	0.017	*****
	SE	0.000069	0.00013	0.007	
	B-std	0.16	0.21	0.090	
AMD*Auditory (AMD*AF)	B	*****	*****	−0.13	−0.059
	SE			0.010	0.012
	B-std			−0.23	−0.088
Constant					
	B	−0.0012	0.0067	10.26	3.81
	SE	0.0013	0.0023	0.15	0.099
	p	0.38	0.004	<0.001	<0.001

The terms, AF and AMD*ISp were not included as predictors in any of the models, and thus not included in this table. [*****] designates terms not included in a given model.

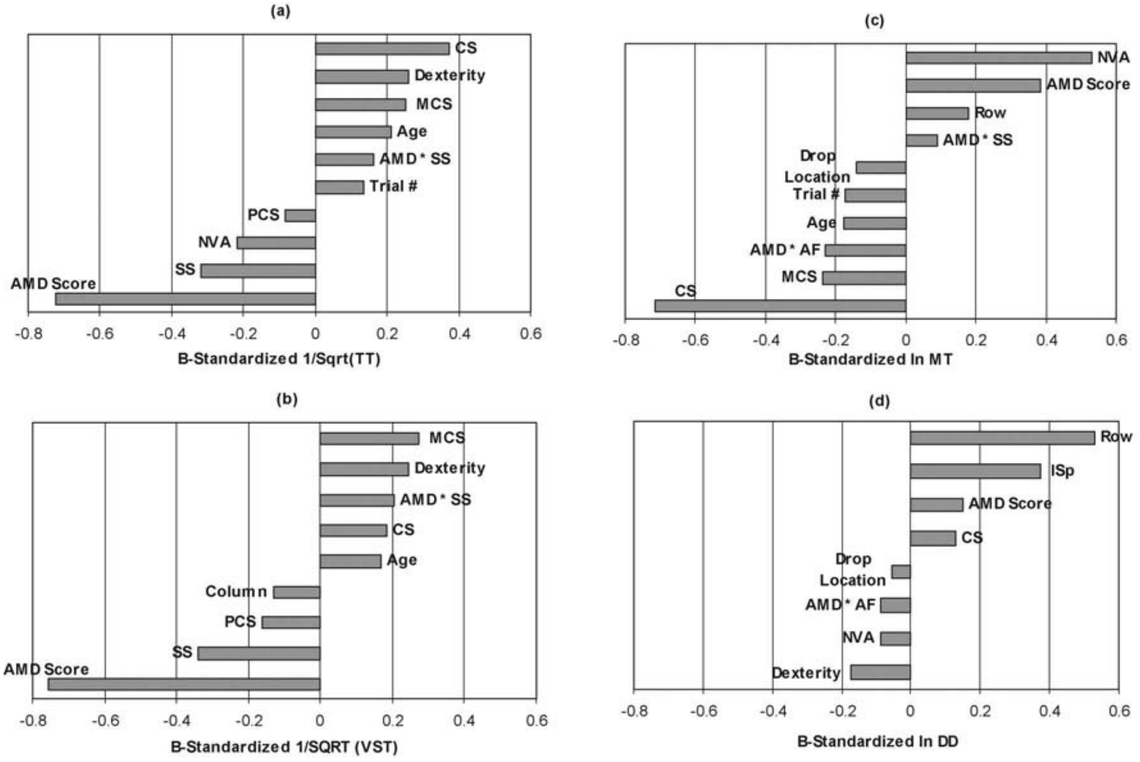

Figure 3 (a–d). Relative impact of predictor variables, B-std.

Table 4. Transcription of participant verbalized likes and dislikes about the task on the handheld.

Likes

The ability to judge and distinguish cards
(I) appreciated that I got to play a game
I liked the challenge of it
To me, it was fun
Moving the diamonds and hearts, because you can't miss those

Dislikes

I didn't like when the sound was absent
Not really my thing, but didn't dislike anything in particular
It went kind of fast ... the pace was fast
The very minute numbers and images
I would have rather played a real game

participants related the task to playing a game, and made comments about their strategies, or how they liked the challenge. The participant who stated his appreciation of being able to play a game later informed the experimenters that he had regularly played solitaire on the computer until his vision had deteriorated in the past year. One exception was a control participant, who commented that he would have rather played a more challenging game. One participant, under dislikes, asserted his aversion to the trials in which the auditory feedback was absent, thus suggesting his perceived utility for the auditory sound.

In terms of their perceptions of their own performance and their interactions with the task, participants rated their: (1) overall performance; (2) perceived difficulty of the task; (3) perception of how much effort was put forth to complete the task; and (4) perception of the frustration experienced. These questions provided 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).

Figures 4(a–d) illustrate participants' responses to questions of workload. Each workload factor was rated by the participants on a scale from 0 to 10. For performance, a score of 0 means the lowest or worst possible perceived performance level, and 10 the absolute best performance. For Difficulty, a score of 0 equates to no difficulty encountered and 10 means the maximum level of difficulty experienced. For Effort, 0 signifies no effort and a value of 10 translates to the maximum amount of effort applied to complete the task. Finally, in terms of Frustration, 0 means no frustration, and 10 means the maximum amount of frustration was experienced during the course of task completion.

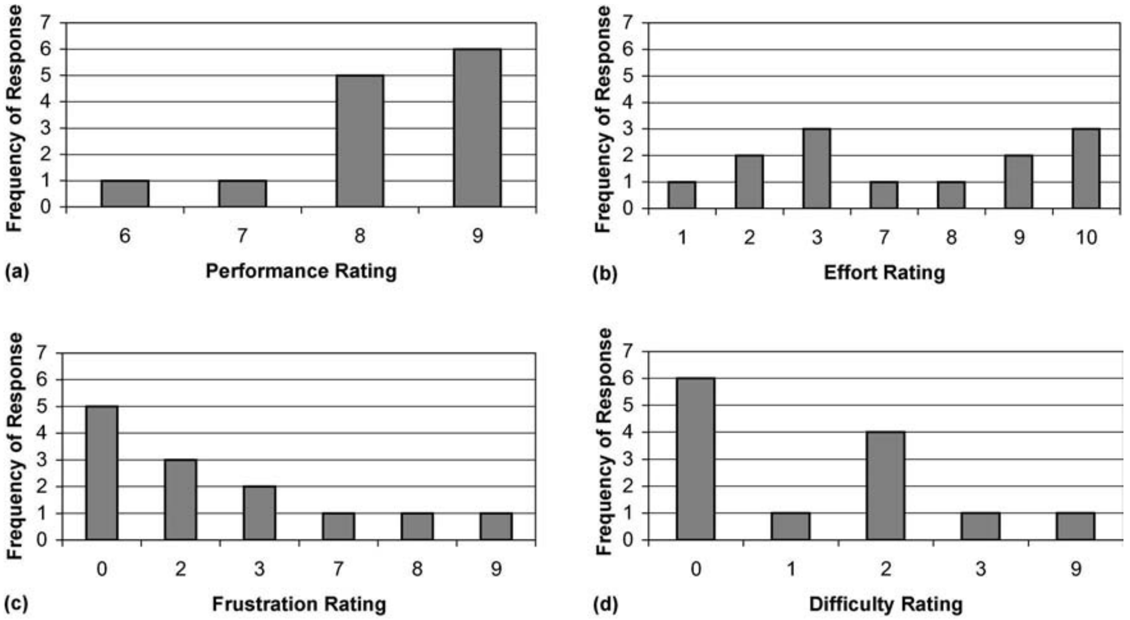

Figure 4 (a – d). Summary of responses to perceived workload subscales.

All participants rated his or her overall performance on the task with a score of 6 or better, with the majority scoring their performance an 8 or 9. 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 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.

In addition to participants' perceptions of the handheld computer technology and their own performance, the exit survey also captured their opinion of the sounds and perceived helpfulness of this supplemental auditory cue. Interestingly, participants' reactions to the sound were consistent with its impact on performance. Responses were a mixture of both positive and negative, just as the auditory cue was not helpful across all participants. In expressing their comfort level with the sound (very comfortable, comfortable, neither comfortable nor uncomfortable, uncomfortable, or very uncomfortable), 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 reporting neither uncomfortable nor comfortable.

Participants' responses to the question: *How helpful was the sound to your completion of the task?* (*very helpful, helpful, neither helpful nor unhelpful, unhelpful,* or *very unhelpful/distracting*) resulted in an even more varied set responses. Three of the participants rated the sound as very helpful, two rated the sound as helpful, 6 of the 13 responded that it was neither helpful nor unhelpful, one rated it as unhelpful, and the final participant rated the sound as very unhelpful/distracting. Figure 5 summarizes these responses. In addition, the participants provided free responses about their general thoughts regarding the sound, which appear in table 5, organized by positive, negative and mixed opinions.

Further analyses were applied to the 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 used 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 (e.g. those with the worst functional vision). 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 on AMD Score (disease severity) and perceived visual functioning (using the VFQ-25).

Briefly, the VFQ comprised a series of 25 questions that queried participants on the degree to which their visual

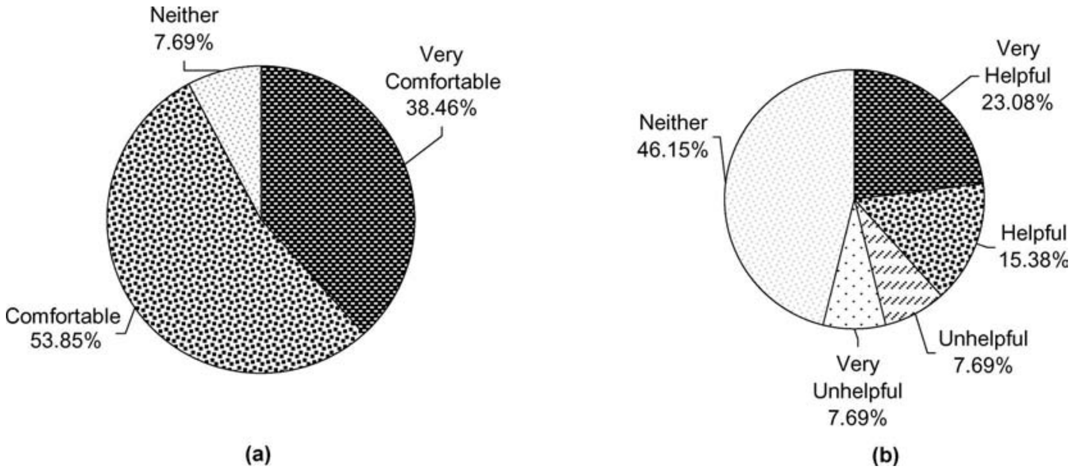

Figure 5. Summary of participant response to questions regarding (a) their comfort level with the auditory feedback and (b) their perception of the auditory feedback helpfulness.

Table 5. 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

impairment interfered with activities of daily living (e.g. near vision activities, driving, mental health, etc.). Scores were tabulated (according to methods used by Mangione *et al.* 2001) and can range from 0 to 100, where 100 is associated with a high quality of life with respect to visual function. For this study, the VFQ scores of the handheld experiment participants ranged from 47.5 to 99.03, mean = 82.83; median = 88.20. Because the data did not meet the assumptions required for the application of parametric statistics (and transformations proved inconclusive), comparisons were made using the Mann–Whitney test for non-parametric comparisons on AMD Score and VFQ for perceived helpfulness of auditory feedback. Figures 6(a) and (b) illustrate the differences between the perceived helpfulness on AMD severity score (figure 6(a)) and VFQ (figure 6(b)). Results revealed a significant difference of VFQ overall between the two groups (Mann–Whitney $U = 5.5$, $Z = -2.13$, $p = 0.034$), and not on the AMD severity score ($p = .05$). In terms of the VFQ, results demonstrated that participants 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 more severely impaired due to their declining visual function

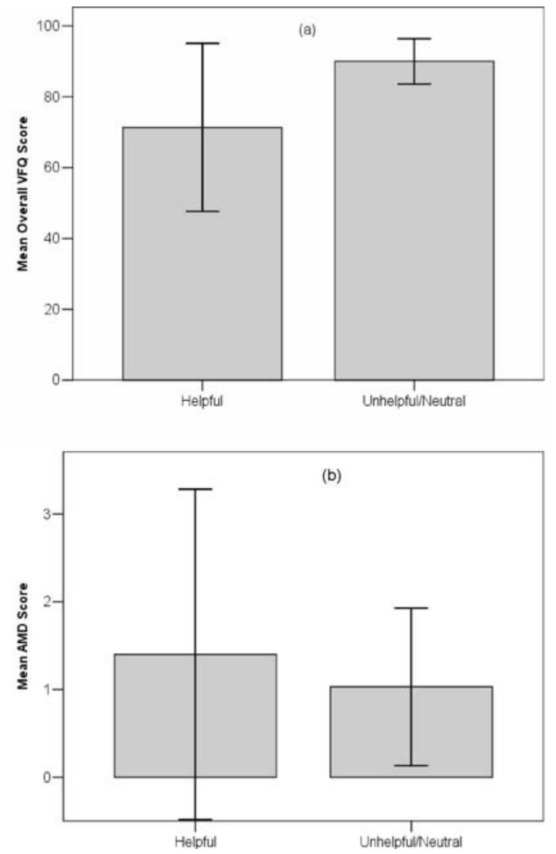

Figure 6 (a, b). Comparisons of, VFQ and AMD Score between participants based on their perception of auditory feedback helpfulness.

are more willing to consider the efficacy of non-traditional, non-visual supplemental cues to integrate with their residual vision.

4. Discussion

The results from the regression, the emergent patterns depicted in figures 3(a–d), and the participants' responses to the exit survey provide an insight into some of the most influential factors affecting handheld computer use for individuals with AMD. The results from this study serve as a baseline for future empirical research, and can inspire development of priorities for further investigations. This discussion emphasizes the ways in which results involving human–computer interactions with handheld computers are consistent with previous research involving desktop systems, while also emphasizing new, emergent interaction models unique to handheld devices for this population of users. The following section enumerates the significant outcomes resulting from this study.

4.1 *Ocular health and function*

Outcome #1. *The persistent impact of clinically-derived ocular measures on performance validates several previous studies (Jacko 1999, 2000, Edwards* et al. *2005, Jacko* et al. *2005) and effectively extends the theory to new interaction platforms with small visual displays.*

Based on the standardized coefficients, measures of visual function dominated performance on visual search, icon movement, and trial time. AMD Score and Contrast Sensitivity (CS) were reliable and dominant predictors of performance in the models all four measures (the only two predictors to be included in all four). This reaffirms the importance of investigations that focus on the sizable impact of visual dysfunction on GUI-based tasks across platforms. More specifically, these models enable the assessment of productivity costs incurred by this population with the handheld computer.

Outcome #2. *Design efforts and strategies aimed at assisting visual search are an appropriate starting point for the development of accessibility solutions for handheld human–computer interactions.*

As the severity of disease (measured by AMD score) worsened (the value increased), all the models reflected performance decrements. The performance differential imposed by AMD score is consistent with the findings of Jacko and colleagues (2005), who observed a similar effect of disease on performance of a simple drag and drop (single file to single folder and no distracters) between a cohort with AMD and visual healthy controls. The AMD score had its most notable influence on TT and VST, and was the third most influential factor on MT. The importance of TT conveys the measurable performance differential incurred due to visual dysfunction on overall task completion, while the magnitude of influence on VST suggests that supporting visual search is an essential component in the quest for accessible design for this population.

Outcome #3. *The role of contrast sensitivity as an essential determinant of task performance for people with visual impairments extends from traditional desktop environments to mobile device use.*

Changes to contrast sensitivity (CS) systematically impacted the efficiency of the task. Improvement (increase) in contrast sensitivity scores emerged as a predictor of faster TT, VST and MT. In previous studies, contrast sensitivity was found to be influential across several desktop computer tasks; the observed influence in this model extends this phenomenon across mobile interaction platforms (Emery *et al.* 2001, Jacko *et al.* 2002a, Edwards *et al.* 2005).

Near visual acuity (NVA), while not influential on the prediction of VST, was included as a predictor in the models of TT, MT and DD. As near visual acuity degraded (increased in value), both TT and MT were slower which confirms the role of the quality of residual vision on task efficiency.

Outcome #4. *The speed–accuracy (movement time–drag distance) tradeoff observed for handheld human–computer interaction is a function of contrast sensitivity (CS) and near visual acuity (NVA).*

The impact of the visual factors on DD was small in magnitude compared to their more substantial influence on the other three efficiency models. Features of the interface, such as the location of the target (drop location) and the Inter-Icon Spacing (ISp) pose a greater influence than the ocular measures on the model of DD. Even so, the influence of the ocular factors on the icon movement component of the task demonstrated unexpected trends in both MT and DD.

Figures 3(c) and (d) illustrate that deteriorations to contrast sensitivity (decreased CS value), diminished near visual acuity (NVA value increase), and worsened AMD severity (increased AMD Score) all contribute to slower movement times (MT). This monotonic trend is not reflected in the models of DD. The impact of worsened AMD Scores consistently contributes to longer modeled dragging distances (DD), but deterioration to CS and NVA induce improvement (shortened) in drag distance. Ironically, the models suggest that users with more severe dysfunction in CS and NVA drag the card icons via more efficient and accurate paths to the drop pile. Alternatively, these relationships suggest that while participants with better CS and NVA are modeled to be faster with their use of the stylus to move the card to the drop pile (lower MT), they were less accurate with respect to the efficiency of the path taken to the drop pile (higher DD).

Speed–accuracy tradeoffs are a common occurrence in human-integrated systems, and have been observed in several domains that operate on a combination of discrete and continuous motor control (Fitts 1954, Pew 1969). These types of tradeoffs have also been observed in the performance of older adults' control of arm movements (Darling *et al.* 1989), and in direct manipulation on GUIs (Gillan *et al.* 1990). While the emergent tradeoff in the drag and drop is supported by the underlying theories of Fitts' law, it poses an inconsistency with previous investigations of iconic manipulation for similar user populations (Jacko *et al.* 1999, 2003, 2005). In these studies, declines in CS and NVA function were linked with diminished performance across all facets of the GUI manipulation, and no such tradeoffs emerged. When comparing these investigations with the present study, the most salient variation is the platform considered.

The small display size of the handheld is therefore likely to be driving this effect. The nature of the visual stimulus on the handheld device poses demands on the visual sensory function very different from those associated with the desktop, primarily with respect to the area within which a user performs visual scanning and tracks icon movement. The small display size enables those individuals with higher levels of CS and NVA to take longer paths to the drop pile without imposing extra time to complete their interaction. These users have the ability to dynamically track the icon as they move it (quickly) in non-optimal paths to the destination without sacrificing overall task efficiency. Likewise, those users experiencing lower levels of CS and NVA must be more attentive to visually track the movement of the icon on the display, moving the icon more slowly, and avoiding deviations in the path, as it is more difficult for them to discern the icon from the distracters amid the background, if they lose sight of it in their residual vision.

Figure 7 simulates the impact of moderately severe AMD on the task for both the handheld and desktop displays, at the same viewing distance. 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 device. The effects of increased AMD severity are amplified on the handheld display, as observed in the analyses of the efficiency and accuracy outcome measures. This demonstrates the implications of a small display on performance for users who have AMD. The lower bandwidth of visual information available to users with AMD is further reduced in the presence of the smaller visual display. The entire handheld display could be effectively occluded by central visual field loss. As AMD increases in severity, the central vision is more severely impacted, and the eye and head movements required to target the display within the useful residual vision would take longer. In comparison, on a desktop computer,

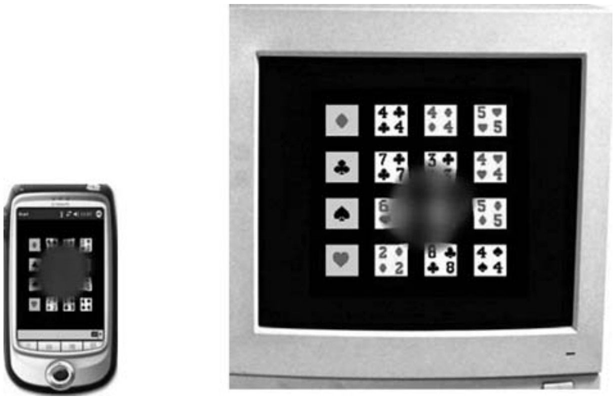

Figure 7. Illustration of the relative impact of AMD on handheld versus desktop, for moderate stages of AMD visual field disruption consistent between the desktop and handheld image (displays not to scale) (Image simulated from MD Simulator 3.1, www.opticaldiagnostics.com).

the larger display will likely subtend the central visual field of all users, requiring all users to exercise a more involved visual search strategy including head and eye movements that lead to additional time on the task.

Despite the measured effects of AMD severity level on the handheld device, 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 reduced ocular function. 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.

4.2 *Personal traits*

***Outcome** #5. Personal characteristics such as age (and thus, indirectly, task familiarity), dexterity, and learning are all influential factors in handheld device use and should be considered in empirical studies involving older adults with visual impairments who are tasked with using handheld devices.*

A handful of personal characteristics proved influential on the models across the different task phases. Increases in age were included as predictors of faster TT, VST and MT, which is contrary to observations obtained in previous studies, where older age was a significant predictor of longer task completion times (Edwards *et al.* 2005). This result can be attributed to the choice of playing cards as the visual icons in the interface. 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 participants were not yet retired). The use of familiar icons can increase users' comfort levels and proficiency with new technologies, which should be explored in future studies. Also, it should be noted that the study of Edwards *et al.* (2005) focused on diabetic retinopathy, a disease affecting a greater range in age. That said, our results provide explicit insight into the older adult population, and how 'young-old' (50–65) individuals differ from those considered part of the 'old-old' segment.

Dexterity was found to be influential in models of TT, VST and DD. As dexterity improved (or the score increased) TT and VST were faster, and DD was shorter, indicating a more efficient interaction. Over time, additional fatigue could amplify the impact of dexterity, especially for older adults. Also, the selection of input device is a feature of the interface that is easily altered to accommodate a range of individual needs. The implications of input device on a small interactive display are critical to the successful interaction and thus the small relative magnitude of this effect should not be overlooked. This result also suggests that dexterity is linked to visual search, and implies the use of the stylus as a pointing mechanism to direct visual search.

The impact of Trial # indicated that participants demonstrated faster MT and TT during later trials, a small learning effect. Interestingly, this effect was not realized for VST, suggesting that for this task and set of participants, practice improved control of the stylus, but not the ability to locate the icon over time. The lack of a practice effect on efficient visual scan is again likely to be linked to the small size of the display. This suggests that participants were able to improve their interactions with the stylus over time, while their times for searching for the icon did not incrementally improve. Again, this is another example that older adults with visual impairments are wholly capable of the interactions required by a small handheld computer. However, unless their ability to extract visual information from the display is supported through strategic design, the potential applications for handheld computers will not be usable.

4.3 *Interface characteristics*

***Outcome** #6. Older adults with visual impairments are able to use a stylus for input on a handheld device and the ease with which the stylus is operated influences several key aspects of interaction.*

Column (the column where the target icon was located, per trial—leftmost, middle, rightmost) impacted VST, while the Row where the target card icon was located in each trial impacted the MT and DD. Columns further to the right yielded increased VST, consistent with the nature of visual scan for Western users, who work from left to right to locate an icon. The impact of rows lower on the display also increased DD and MT, suggesting that participants had more difficulty making use of the stylus to move icons from lower sections of the display. In addition, as the location of the drop pile moved lower on the display, the predicted MT and DD also increased. This is likely to be related to the ease with which the participants operated the stylus. It is surprising that, even though the display on the handheld spanned the visual field, there is still measurable complexity in the identification and tracking of the icons across the display.

***Outcome** #7. Supplemental non-visual cues may prove valuable in making handheld devices more accessible to individuals with visual impairments.*

The main effects of Auditory Feedback (AF) were not included in any predictive models. However, the AF*AMD interaction was influential in decreasing MT and DD. This is especially important in light of the considerable negative impact that ocular disease and functional impairment imposed on the performance models, as discussed previously. The performance gains realized from the inclusion of auditory feedback increased as AMD severity level worsened. The effect on DD suggests that participants with

AMD spent more time 'chasing' the drop pile, in the absence of auditory cues. This implies that supplemental non-visual cues may prove valuable in facilitating accessibility to these devices. However, based on comparisons of the standardized coefficients for AMD score and AMD*AF, there remains room for improvement to more fully counteract the impact of disease on the release of the icon.

Outcome #8. *Auditory feedback demonstrates utility as a solution to improved access, for individuals with visual impairment.*

The conclusions concerning auditory feedback are consistent with the findings of Jacko and colleagues in their examination of the drag and drop (Jacko *et al.* 2005). The authors found that individuals with the most severe visual dysfunction experienced the most significant gains in performance with the inclusion of supplemental non-visual cues. Also consistent is the fact that the presence of auditory feedback did not degrade the performance of those without ocular pathology, supporting its utility as a universal solution to improved access. Furthermore, the qualitative information collected from these participants suggests that individuals who perceive their visual impairment as interfering with their activities of daily living were more likely to identify an assistive intervention as helpful to their use of the technology. Auditory feedback poses a low-cost, easily implemented means to counteract performance decrements imposed by visual dysfunction.

Outcome #9. *Consistent with traditional desktop displays, older adults with visual impairments using handheld devices also experience difficulties tracking target icons amongst distracters present on the display.*

It is intuitive that SS was found to be influential on VST in that it sufficiently imposed predicted increases in TT, slowing the rate of task completion. In addition, the SS*AMD interaction was found to have a significant influence on TT, VST, and MT. There was a predicted increase in MT as a result of the SS*AMD interaction, indicative of participants' difficulties with tracking an icon amongst an increasing number of distracters across the display. However, the influence of SS*AMD on visual search was not intuitive. The model suggests that while increased SS imposes longer TT and VST, in the presence of AMD, the impact of set size is subdued. In fact, at the most severe levels of AMD, the effects of increased sets size are entirely canceled out.

This is attributed to the small display size. The entire display of the handheld device is easily viewed within the central visual field. For an individual with severe visual impairments, the number of icons at the onset of the trial display may appear constant. This is depicted in figure 8. Visually healthy users are likely to be changing their visual search strategies based on the number of icons that appear at display onset. The users with AMD, however, are not changing their visual search strategies in the presence of different numbers of distracter icons, because with an occluded visual field, a display with 4 card icons may initially appear the same as a display with 12. This outcome is very relevant to design, as users with impairments may not perceive changes to the display as quickly as their visually healthy counterparts. If presenting critical information, designers should consider incorporating auditory cues in order to direct the user's attention.

Outcome #10. *Design theories for traditional desktop environments should not be applied automatically to alternative platforms such as handhelds.*

Design guidelines and principles for desktop design were not purely reflected in the models of handheld human–computer interaction for this population. Specifically, the effects of diminished spacing did not influence longer search and selection times, as demonstrated in research on computational HCI models (of desktop interactions) for a visual healthy population (Hornoff 2001, Everett and Byrne 2004). Both studies found icon strategy to deviate as a function of the proximity of the distracters. Everett and Byrne observed inefficient search patterns when the icons were in close proximity. Likewise, Hornoff observed users exhibiting slower but more accurate icon search and selection interactions when icons were in close proximity.

The absence of interesting effects due to spacing for the present task and user population is likely due to the spacing size relative to the physical display size of the handheld device, and because the effects of visual health were more prevalent. Regardless, it suggests that existing computational models for HCI need to be reassessed prior to their application to alternative GUI platforms, such handheld computers, and for alternative users groups.

5. Conclusions

The most compelling outcome from this study is its demonstration that this group of older adults, both with and without visual impairments, were capable of the successful interaction required to interact with non-traditional IT platforms, namely handheld computers. To reiterate, all participants demonstrated high levels of task accuracy, completing 97% of trials involving dropping the correct card into the correct pile. However, the efficiency of the task was compromised largely in the presence of increased visual dysfunction and disease severity.

The regression models demonstrated the potential for low-cost, easily implemented design interventions (e.g. auditory feedback and increased display contrast) to enhance task efficiency for older individuals with visual impairments to levels equivalent to those of users without ocular pathology. This study presents a strong argument in

Figure 8. Illustration of set size effect for participants diagnosed with AMD (Image simulated from MD Simulator 3.1, www.opticaldiagnostics.com).

favor of continued research in the area of mobile computing for these population segments, as the interactions and strategies can deviate from those traditionally observed in the context of desktop computers. Based on the outcomes from this study, the performance decrements incurred by visual search are some of the most difficult to overcome, even with practice.

While the participant population utilized in this study cannot conclusively account for the heterogeneity in mental and physical capabilities of the aging population, it strongly evokes a need to further explore handheld human–computer interaction with the aging population. This is especially true in light of the fact that in the near future, the older adult population will be comprised of baby boomers, who, like the participants in this study, will have at least some regular experience with the use of GUIs on a range of information technologies. The demand for usable, useful mobile technologies, with small displays akin to the handheld computer, will no doubt rise proportionately to the rapid growth in the 65 years and older population segment.

The unique relationships revealed for handheld human–computer interaction demonstrate the additional dimension that small changes in hardware can impose on the assessment of interactions. While existing research that reports desktop HCI contributes a useful baseline for the assessment of interactions, characteristics of the new platform, such as input mechanism, size, and context, must be taken into account. Future work will consider the effects of varied contexts of use, additional critical GUI manipulations, as well as different types and levels of visual dysfunction for HCI with small mobile displays, using the outcomes from this study as baseline research. Given the high aptitude of this user population in terms of task completion, the impact of time constraints on task accuracy will be applied in future work. In addition, a comprehensive comparison between desktop and handheld HCI could facilitate an improved understanding of how to apply existing design guidelines to a handheld computing environment. Controlled empirical research investigations, similar to what has been presented here, are necessary to build a framework of interaction thresholds that can anticipate the wide variety of user needs as new interaction paradigms are introduced by emerging technologies. This serves to bridge, as opposed to widen, the digital divide for users with divergent needs for uninterrupted access to the benefits of emergent information technologies.

Acknowledgements

This research comprised a portion of V. Kathlene Leonard's dissertation and was made possible through funding awarded to Julie A. Jacko by the Intel Corporation and the National Science Foundation (BES-9896304). Young Sang Choi, Kevin Moloney, and Ji Soo Yi are graciously acknowledged for their assistance in carrying out the experimental protocol. We also acknowledge the support of Nova Southeastern University, who conducted patients' ocular examinations and for providing the facilities within which the experimentation was conducted.

References

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

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

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

Barnard, L., Yi, J.S., Jacko, J.A. and Sears, A., 2005, An empirical comparison of use-in-motion evaluation scenarios for mobile computing devices. *International Journal of Human–Computer Studies*, **62**(4), pp. 487–520.

Barnard, L., Yi, J.S., Jacko, J.A. and Sears, A., in press, A new perspective on mobile device evaluation methods. *Personal and Ubiquitous Computing*.

Bird, A.C., Bressler, N.M., Bressler, S.B., Chisolm, I.H., Coscas, G., Davis, M.D., Jong, P.T.V.M.D., Klaver, C.C.W., Klein, B.E.K., Klein, R., Mitchell, P., Sarks, J.P., Sarks, S.H., Soubrane, G., Taylor, H.R. and Vingerling, J.R., 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), pp. 367–374.

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

Brewster, S., 1998, Sonically-enhanced drag and drop. *International Conference on Auditory Display* 1998, Glasgow, pp. 1–7.

Coyle, J.T., 2003, Use it or lose it—do effortful mental activities protect against dementia? *New England Journal of Medicine*, **348,** pp. 2489–2490.

Craven, J., 2003, Access to electronic resources by visually impaired people. *Information Research*, **8**(4), paper number 156. Available online at: http://information.net/ir/8-4/paper156.html.

Darling, W.G., Cooke, J.D. and Brown, S.H., 1989, Control of simple arm movements in elderly humans. *Neurobiology of Aging*, **10,** pp. 149–157.

Edwards, P.J., Barnard, L., Emery, V.K., Yi, J.S., Moloney, K.P., Kongnakorn, T., Jacko, J.A., Sainfort, F., Oliver, P.R., Pizzimenti, J., Bade, A., Fecho, G. and Shallo-Hoffmann, J., 2004, Strategic design for users with diabetic retinopathy: factors influencing performance in a menu-selection task. In *The Sixth International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS 2004)*, Atlanta.

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

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

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

Everett, S.P. and Byrne, M.D., 2004, Unintended effects: varying icon spacing changes users' visual search strategy. *CHI Letters*, **6**(1), pp. 695–702.

Fahrmeir, L. and Tutz, G., 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,** pp. 381–391.

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

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

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

Hart, S.G. and Staveland, L.E., 1988, Development of NASA-TLX (Task Load Index): results of empirical and theoretical research. In *Human Mental Workload*, N. Meshkati (Ed.), pp. 139–183 (Amsterdam, The Netherlands: Elsevier Science Publishers B. V. (North-Holland)).

Hornoff, A., 2001, Visual search and mouse-pointing in two-dimensional visual hierarchies. *ACM Transactions on Computer–Human Interaction*, **8**(3), pp. 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), pp. 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), pp. 151–164.

Jacko, J.A. and Leonard V.K., in press, Satisfying divergent needs in user-centered computing: Accounting for varied levels of visual function. *Human Factors and Ergonomics Society, Annual Review of Human Factors and Ergonomic Subcommittee*, R. Williges (Ed.).

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

Jacko, J.A., Barreto, A., Scott, I.U., Rosa, R.H. and Pappas, C.J., 2000a, Using electroencephalogram to investigate stages of visual search in visually impaired computer users: preattention and focal attention. *International Journal of Human–Computer Interaction*, **12**(1), pp. 135–150.

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

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

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

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

Jacko, J.A., Scott, I.U., Sainfort, F., Moloney, K.P., Kongnakorn, T., Zorich, B. and Emery, V.K., 2002b, Effects of multimodal feedback on the performance of older adults with normal and impaired vision. *Lecture Notes in Computer Science*, **2615,** pp. 3–22.

Jacko, J.A., Scott, I.U., Sainfort, F., Barnard, L., Edwards, P.J., Emery, V.K., Kongnakorn, T., Moloney, K.P. and Zorich, B., 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), pp. 33–40.

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

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

Jacko, J.A., Moloney, K.P., Kongnakorn, T., Barnard, L., Edwards, P.J., Emery, V.K., Sainfort, F. and Scott, I.U., 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**(2), pp. 183–218.

Jacko, J.A., Barnard, L., Yi, J.S., Edwards, P.J., Leonard, V.K., Kongnakorn, T., Moloney, K.P. and Sainfort, F., 2005, Empirical validation of the Windows accessibility settings and multimodal feedback for a menu selection task for users with diabetic retinopathy. *Behaviour & Information Technology*, **24**(6), pp. 419–434.

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

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

Macular Degeneration Partnership. What is AMD? Available online at: http://www.macd.net/amd/index.asp (accessed 5 November 2002).

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

MD Foundation, 2004, Pictorials descriptions of AMD. Available online at: http://www.eyesight.org/Pictorials/Pic-Look_Like/pic-look_like.html (accessed 29 July 2004).

NASA Ames Research Center, 1987, NASA Human Performance Research Group Task Load Index (NASA-TLX) instruction manual. Nasa Ames Research Center.

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

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

Pew, R.W., 1969, The speed-accuracy operating characteristic. *Acta Psychologica*, **30,** pp. 16–26.

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

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

Scott, I.U., Feuer, W.J. and Jacko, J.A., 2002a, 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), pp. 850–857.

Scott, I.U., Feuer, W.J. and Jacko, J.A., 2002b, 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), pp. 857–862.

Sears, A., Lin, M., Jacko, J. and Xiao, Y., 2003, When computers fade. Pervasive computing and situationally-induced impairments and disabilities. In *Human–Computer Interaction: Theory and Practice (Part II)*, C. Stephanidis and J. Jacko (Eds), pp. 1298–1302 (Mahwa, NJ: Lawrence Erlbaum Associates).

Tiffin, J. and Asher, E.J., 1948, The Purdue Pegboard: norms and studies of reliability and validity. *Journal of Applied Psychology*, **32,** pp. 234–247.

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

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

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