Usability Evaluation of Notebook Computers and Cellular Telephones Among Users With Visual and Upper Extremity Disabilities

Aaron M. Mooney

Thesis submitted to the Faculty of the Virginia Polytechnic Institute

and State University in partial fulfillment of the requirements for the degree of

Master of Science

in

Industrial and Systems Engineering

Dr. Maury A. Nussbaum, Co-Chair

Dr. Tonya L. Smith-Jackson, Co-Chair

Dr. H. Rex Hartson

July 8, 2002

Blacksburg, Virginia

Keywords: Visual disabilities, physical disabilities, usability, notebook computers, cellular telephones, critical incident technique, remote-ethnographic method, universal design, design guidelines, special needs.

This page intentionally left blank for double-sided printing.

Usability Evaluation of Notebook Computers and Cellular Telephones Among Users With Visual and Upper Extremity Disabilities

Aaron M. Mooney

(Abstract)

Information appliances such as notebook computers and cellular telephones are becoming integral to the lives of many. These devices facilitate a variety of communication tasks, and are used for employment, education, and entertainment. Those with disabilities, however, have limited access to these devices, due in part to product designs that do not consider their special needs. A usability evaluation can help identify the needs and difficulties those with disabilities have when using a product and universal design principles can then be applied to enhance accessibility and usability. This study addresses the usability of two of the most common information appliances – notebook computers and cellular telephones.

The usability of notebook computers was evaluated using a remote ethnographic method where participants recorded usability-related critical incidents. Participants included those with a wide range of abilities, such as legal blindness, total blindness, and upper extremity physical disabilities. Objective and subjective measures were used to determine the effects of several specific design parameters for cellular telephones.

The notebook computer study revealed that participants have difficulty with non-standard keyboard layouts, the use of isometric pointing devices, case latches, and inadequate system feedback. User performance and ratings in the cellular telephone study were the best with the 12 mm lateral pitch and 0.7 mm key height, while the fewest task failures were committed using the 0.5 mm keystroke. Participants also preferred telephone models with large <Power>, <Send> and <End> keys located in prominent locations, and 22-point and 36-point display fonts. These results were used to generate product-specific design guidelines that can be used to design notebook computers and cellular telephones that are more usable and accessible for users with visual and upper extremity physical disabilities. Universal design implications are also discussed.

Acknowledgements

I would like to thank Dr. Maury Nussbaum for his input, review, and support for this thesis. Dr. Nussbaum was always readily available to review my work, provide suggestions, and just listen and when frustrations arose. Additionally, I appreciate his efforts to quickly turn around my work so I could continuously improve the quality of the work, yet keep deliverables on schedule.

I would also like to thank Dr. Tonya Smith-Jackson for her help in beginning the research process, as well as her help and support throughout. Dr. Smith-Jackson helped with participant recruiting, reviewed the work, and provided many suggestions on improvements when necessary. Additionally, I would like to thank Dr. H. Rex Hartson for his flexibility and willingness to help in a pinch, and Lovedia Cole for her knowledge of graduate school procedures, cheery attitude, and willingness to help students navigate the myriad of nebulous procedures involved with graduation.

Many thanks to Jeff Lancaster, Bill Penhallegon and Chuck Perala for their friendship. Life in Blacksburg, and life in graduate school would not have been nearly as enjoyable without your friendship.

Finally, I would like to thank Shannon Green for her love, friendship, and emotional support while I completed this program. Shannon was always willing to listen and provide words of encouragement when needed. Without her love, friendship, and support, this work would not have been possible.

Table of Contents

List of Tables

List of Figures

1.0 - Introduction

Information appliances such as cellular telephones and notebook computers are becoming integral to the lives of many, and facilitate rapid access to communication and information services (Vanderheiden and Henry, 2001). These devices are becoming more common for activities such as work, communication, education, and entertainment (Vanderheiden and Henry, 2001). The decreasing cost and increased availability of these devices is helping to increase their proliferation and use within society. Due to the design of these devices, however, use is not proliferating among those with disabilities, resulting in limited access to important information and communication services (Kaye, 2000).

The evolution of technology within society has shifted the economic paradigm from the production of physical goods to the exchange of information (Stephanidis, 2001). As this trend continues, more telecommunications and information services will emerge that require the use of information appliances (Stephanidis, 2001). Contributing to this paradigm shift are advances in technology, continued decreases in the size and price of technology, and increased societal pressure to use information appliances (Bass, 2001).

The ability to quickly and conveniently access information is important within an information society (Bass, 2001). Competitive pressures within the business environment often require rapid decision-making. Rapid and convenient access to information helps many workers communicate and make important decisions in a timely manner. Information appliances such as cellular telephones and notebook computers facilitate this communication by allowing workers to receive important telephone calls and electronic mail (e-mail) at any time, from almost any location (Bass, 2001). Additionally, the ability to use the Internet to acquire information at any time, from nearly any location can assist workers in making informed decisions in a timely fashion.

Those in the business environment, however, are not the sole beneficiaries of the use of these information appliances. Notebook computers and cellular telephones are often used for personal needs. Families can use notebook computers to send e-mail to distant relatives, to look up information while on a family vacation, or to get information on local restaurants and entertainment while traveling. Cellular telephones can be used to coordinate activities between spouses, receive directions when driving, locate errant children, or even call for emergency assistance.

Personal use of these devices is not limited to family-related activities. According to Bikson and Panis (1995), individuals that use technology-related devices can have an increased knowledge of current events, and a higher affiliation with others that share their interests, compared to those that do not use these devices. Facilitation of communication through the use of e-mail, chat groups, and discussion boards can help users affiliate with others that share common interests. For those that live in remote areas, lack socials skills, or lack mobility, this affiliation may become an important social and affective support structure (Bikson and Panis, 1995; Pieper, 2001).

Educational institutions are becoming more reliant on the use of information appliances to accentuate and sometimes even present instruction. Courses at Universities such as Virginia Tech, require students to access class materials and submit assignments via the internet. Virginia Tech is increasing its commitment to the ready access of information and the use of information appliances, by wiring new buildings with an ample supply of network connections and providing computers in the classroom to present instructional material. Classrooms in older buildings will be retrofitted with wireless Internet technology to allow for the use of information appliances (*College of Engineering Computing Requirements*, 2001).

In addition to the commitment to increase the use of information appliances through infrastructure improvements, educational institutions are determined to increase their use via department requirements. In the Fall of 2002, new students at Virginia Tech in the Pamplin College of Business, will be required to own a notebook computer and The College of Engineering is contemplating a similar requirement (*Pamplin College of Business Student Computer Purchase Policy Statement*, 2001; *College of Engineering Computing Requirements*, 2001). Due to these upgrades in network infrastructure, and the requirement for students to own and use notebook computers, it can be anticipated that computers may be used to facilitate instruction in many classes.

The access to these devices, and the benefits that this access provides, is not available to all members of society. According to Kay (2001), those with disabilities may only be half as likely to access and use this technology as those without disabilities. This disparity may be due in part to the limited accessibility of these devices (Kay, 2001). Benyon and Crerar (2001) define accessibility as the precursor to usability and the ability to "physically access equipment" as well as the "operational suitability of both hardware and software for *any* potential user". In some cases, a product may be marginally accessible though not usable. Before a product can be deemed "usable" it must first be "accessible" (Benyon and Crerar, 2001). Without the ability to access and use information appliances, those with disabilities may be restricted from the business, personal, and education benefits that this access to information provides (Emiliani, 2001).

The use of information appliances is important for those with disabilities. Those with disabilities have traditionally had lower levels of employment, and higher levels of poverty when compared to others within society (McNiel, 1997). The easy access and use of information through the use of information appliances can increase access to employment by allowing those with disabilities to use their cognitive skills, rather than physical skills (Langton and Ramseur, 2001).

The use of commonly available information appliances may facilitate employment by other means as well. Lupton and Seymour (2000) indicate that employers may be reluctant to hire those with disabilities due to unfamiliarity with assistive technology. This reluctance may be due to fears that assistive devices will be expensive, unreliable, or require a change in business processes to accommodate the person and the device. The ability to access information using readily available devices may eliminate those fears. Even those with limited mobility would have greater access to employment through the use of these devices, since information appliances may allow workers to complete work assignments and communicate with others from home (Williges and Williges, 1995).

According to Soares and Kirk (2000), to design information appliances that are more accessible, designers must understand the needs, requirements, and preferences of users with disabilities. To understand these issues, designers must meet with users with disabilities to discuss their needs and preferences, and study how they would use these devices if they had access to them. Soares and Kirk (2000) believe that although this sounds simplistic, it is not simple to do.

Newell and Gregor (1997) noted that it may be difficult for designers to meet with those with disabilities to discuss their needs. Depending on the demographics of an area, the availability of those with disabilities may be severely limited. They may also be uncomfortable discussing their design needs with others, or unable to meet with designers due to lack of time and transportation. Newell and Gregor (1997) indicated that additional time commitments and transportation logistics are imposed when studying users with disabilities, and these demands may be more than the designer, design budget, or project timeline is willing to accommodate.

In this study, the usability of notebook computers and cellular telephones was evaluated among participants with visual and upper extremity physical disabilities. The notebook computer study used a remote ethnographic method to collect user-reported critical incidents. The cellular telephone study used performance and subjective measures to determine how usability was affected by several specific design parameters. Participants included those with legal blindness, total blindness, minor and severe upper extremity physical disabilities, and those with no apparent disability. Participants with no apparent disability served as a control group.

Product-specific design guidelines to improve the usability of notebook computers and cellular telephones were created based on study results and implications for universal design were discussed. These guidelines can assist designers with future notebook computer and cellular telephone designs, to help ensure these products are more usable for those with visual and upper extremity physical disabilities.

2.0 - Review of Literature

2.1 - Prevalence of Disabilities

The United States Census Bureau estimates that there are 53 million people in the United States with a disability (McNiel, 1997). Of these, it is estimated that 33 million have a disability that can be considered severe. The U.S. Census Bureau defines disability using several criteria. These criteria include the use of a device to aid in mobility, difficulty performing a functional activity such as seeing, hearing, speaking, or grasping small objects, difficulty with an activity of daily living, having a mental or emotional condition, the inability to work outside the home, or receiving federal benefits based on the inability to work. A person meeting at least one of these criteria is considered to have a disability (McNiel, 1997).

According to Vanderheiden (1990), determining the number of individuals with a disability is difficult. Disability estimates depend on the definition of a disability, and the source of the data. Many data sources are available, and data overlap between sources can make estimates difficult. Since sources classify data based on disability, those with more than one disability may be classified more than once (Vanderheiden, 1990). Nonetheless, Newell and Gregor (1997) estimated the distribution of certain disabilities within the United States as:

- 1 in 10 have a hearing impairment, and 1 in 125 are deaf
- 1 in 100 have a visual disability, 1 in 475 have legal blindness, and 1 in 2000 have total blindness
- 1 in 250 are wheelchair users
- 20% of the population has difficulty performing basic physical activities
- 7.5% are unable to walk, lift, read, or hear without help

2.2 - Employment Statistics for Those with Disabilities

According to the *Americans with Disabilities – Household Economic Studies* report released by The United States Census Bureau, approximately 72% of working age Americans with disabilities are unemployed. Among Americans who are mobility impaired, approximately 25% are employed, with employment levels for those with total blindness or severe visual impairments at approximately 31%, and 44%, respectively. (McNiel, 1997).

2.3 - Causes of Disability

People may acquire a disability due to a number of reasons. Disabilities may be present at birth, or appear through the onset of disease (Vanderheiden, 1990). A person may experience a disability due to an acute or chronic event. Accidents, an acute event, can leave a person with a disability. Medical conditions, such as heart attacks and strokes, while chronic in their development, may create an acute episode that results in a disability (Newell and Gregor, 1997). Chronic degenerative bone or nervous system disorders may also result in the long-term diminishment of abilities (Benyon, Crerar, and Wilkinson, 2001).

Even without the occurrence of an acute or chronic medical event, abilities can diminish through the usual processes of aging (Newell and Gregor, 1997). Throughout the aging process, coordination, mobility, sight and hearing may diminish (Benyon et al., 2001). Stephanidis (2001) and Vanderheiden (2001) state that all people are on an "ability continuum", with some starting near the able bodied side, moving along this continuum and inevitably losing abilities through time, accidents or disease (Stephanidis, 2001; Vanderheiden and Henry, 2001).

Demographic trends in the United States indicate a growing elderly population (Newell and Gregor, 1997; Vanderheiden, 1990). Advances in medical technology have increased the average lifespan in the United States and other developed countries. Diseases that were once fatal can now be managed, resulting in an increased population of elderly that can stay alive longer, but with greater levels of disability (Newell and Gregor, 1997).

Newell and Gregor (1997) indicate that although people may experience a disability through accidents, aging, or disease, they do not want to give up the access to information and technology to which they were previously accustomed. It is expected that 1 in 3 Americans will experience a short-term disability due to injury or illness. Those with temporary disabilities still expect to be able to use their information appliances during recovery. According to Newell and Gregor (1997), those that experience a long-term disability, who used a piece of technology to accomplish common tasks in the past, still expect to be able to use it as their abilities diminish.

Although the desire for accessible products has become more visible due to the growing elderly population, and due to those that have used technology in the past and then experience a disability, the desire for accessible products is not new. Advocates for those with disabilities have been working to bring attention to this issue for many years. Congress, faced with growing concern over the accessibility needs of the elderly and those with disabilities, passed the Americans With Disabilities Act, requiring product manufacturers to consider accessibility when designing products (Center for Universal Design, 1997).

2.4 - Disability Legislation

The Americans with Disabilities Act (ADA) was the initial legislation that brought accessibility issues into public focus (Americans with Disabilities Act Technical Assistance Program, 2001). Title 2 of the ADA, passed in 1990, required services and communications products to be accessible to those with disabilities ("The Americans With Disabilities Act," 1990). If the services or communications products were not already compliant, manufacturers had to make “reasonable accommodations” to demonstrate an effort towards compliance. According to Newell and Gregor (1997), reasonable accommodations are small, low-cost modifications that increase access to those with disabilities. Telecommunications companies provided some accommodations for products, such as a volume control and teletypewriter (TTY) compatibility for those with deafness and hearing impairments (Jacobs, 1999).

After the passage of the ADA, the Federal Communications Commission passed the Telecommunications Act of 1996. Within the Act is the provision that all telecommunications equipment be designed to be accessible by those with disabilities (Telecommunications Act of 1996, 1996). Telecommunications equipment includes, but is not limited to, cellular telephones, personal computers, residential telephones, etc. If equipment is not produced that is accessible by those with disabilities, the Act requires that manufacturers make reasonable modifications to equipment to ensure compatibility with special access technologies, such as TTY, Braille readers, etc.

As reasonable accommodations due to the ADA and Telecommunications legislation began appearing, an interesting phenomenon was observed. Those without disabilities were using accessibility features designed for those with disabilities. This phenomenon became known as the "curb-cut" effect (Jacobs, 1999; Newell and Gregor, 1997). One of the first accessibility modifications due to the ADA was the reduction in sidewalk curb height at intersections and crosswalks. This reduction in curb height, known as a curb cut, allowed those in wheelchairs to get on and off the sidewalks with less difficulty. After the sidewalks were modified, bicyclists and those pushing shopping carts and strollers began to use the curb cuts as well (Jacobs, 1999; Newell and Gregor, 1997).

Electronic devices can exhibit a curb-cut effect (Schneiderman, 2000). Televisions equipped with closed captioning decoders were intended to allow those with hearing impairments to watch TV. Those with normal hearing found novel uses for this technology, such as watching TV in noisy environments, or watching when someone is sleeping in the same room (Jacobs, 1999). Telephones adapted due to the ADA and Telecommunications Act also exhibited an electronic curb-cut effect. The volume controls on telephones designed to allow those with hearing impairments communications access, allowed people with normal hearing to adjust the volume to their preference, or to talk in noisy environments (Jacobs, 1999). Computer related examples of electronic curb-cuts include power switches on the front of computers, adjustable keyboards, and user controls for the levels of volume, and display brightness (Schneiderman, 2000). Jacobs (1999) and Newell and Gregor (1997) indicated that when product manufacturers include accessibility features in product design, those without disabilities enjoy the additional benefits that these features allow (Jacobs, 1999; Newell and Gregor, 1997). Through the use of universal design, product manufacturers can design a product that is accessible, and more functional to both those with disabilities, and those without.

2.5 - Universal Design

Universal design is the design of products and environments to be usable by the greatest pool of people as is practical, without the need for adaptations or specialized design (Center for Universal Design, 1997). The principles of universal design can be used in the design of products to ensure that users with the widest range of abilities have access to them. Universal design principles include equitable use, flexibility in use, simple and intuitive use, perceptible information, tolerance for error, low physical effort, and size and space for approach and use (Center for Universal Design, 1997).

The Center for Universal Design at North Carolina State University, and the Trace Center at the University of Wisconsin, state that products that incorporate these principles can be used by nearly anyone, regardless of abilities, without the need for special equipment or special device modes. These products are easy to understand and use, regardless of the user's level of experience, knowledge, skills, or cognitive abilities (Center for Universal Design, 1997). These devices are easy to reach and manipulate regardless of a users size, posture and mobility. They provide different, and redundant methods of sensory feedback to the user, such as auditory, visual, and tactile feedback (Vanderheiden and Henry, 2001). These devices are comfortable to use, and don't require the use of excessive strength or non-neutral postures to operate (Center for Universal Design, 1997). Devices that incorporate universal design principles have a high tolerance for error, allowing errors to be minimized, recoverable, and not catastrophic (Center for Universal Design, 1997). The principles of universal design are summarized below (Table 2.5.1).

Table 2.5.1 - Universal Design Principles
(adapted from the "Principles of Universal Design",Center of Universal Design, North Carolina State University.)

Principle:	Description:	Guidelines:
Equitable Use	The design is useful and marketable to people with diverse abilities.	a. Provide consistent operation for all users. b. Avoid segregating or stigmatizing users. c. Make design appealing to all.
Flexibility in Use	The design accommodates a wide range of individual preferences and abilities.	a. Provide choice in method of use. b. Accommodate use with either hand. c. Facilitate accuracy and precision
Simple and Intuitive Use	Use is easy to understand, regardless of experience, knowledge, language, etc.	a. Eliminate unnecessary complexity. b. Be consistent with user expectations. c. Provide effective prompting and feedback.
Perceptible Information	Design communicates necessary information to the user, regardless of ambient conditions or sensory abilities.	a. Use redundant modes (pictorial, verbal, tactile) for presentation of information. b. Provide adequate information contrast and maximize legibility of essential information. c. Provide compatibility with techniques used by people with sensory limitations.
Tolerance for Error	Design minimizes hazards and adverse consequences of accidental or unintended actions.	a. Minimize hazards and errors. b. Provide fail-safe features and warn of errors. c. Discourage unconscious actions.
Low Physical Effort	Design can be used efficiently and comfortably with a minimum of fatigue.	a. Allow user to maintain a neutral posture. b. Use reasonable operating forces. c. Minimize repetitive actions and physical effort.
Size and Space for Approach and Use	Appropriate size and space is provided for approach, reach, manipulation and use regardless of user's body size, posture, or mobility.	a. Provide clear line of sight to important elements for seated or standing user. b. Make reach comfortable for seated or standing user. c. Accommodate different hand and grip size.

According to Vanderheiden and Henry (2001), universal design is a term that is often misunderstood. At face value, it seems that "universal design" would mean a design that everyone could access and use. Due to the extremely wide range of human abilities, however, it is not always possible or practical to design a product to have universal accessibility or usability. The inability to accommodate everyone, along with other beliefs, may lead designers to abandon universal design altogether (Newell and Gregor, 1997).

There are many reasons why product designers do not incorporate universal design principles. Newell and Gregor (1997) believe that product designers and researchers prefer products that are on the “cutting edge” of technology, and may view accessibility as not intellectually challenging enough and as “charity work”. Some designers may view accessibility as a “fringe” interest, and believe that accessibility refers to specialized rehabilitation-related devices catering to small markets (Newell and Gregor, 1997). Vanderheiden and Tobias (1998) indicate that universal design is not incorporated into product design due to time constraints and the lack of support from upper management, lack of formal processes to include universal design, and inadequate staff resources. According to Shneiderman (2000), designers worry that accommodating users with disabilities may reduce technology to its “lowest common denominator”, and create systems that constrain innovation or are less useful to others. Schneiderman (2000) argues that these worries are really false dilemmas, since the inclusion of universal design requires innovation and can create new technologies, as well as new markets.

The use of products that incorporate universal design can reduce the social stigmas associated with the use of assistive technology products. According to Lupton (2000) those that cannot use commercially available devices must use specialized devices that are expensive and not readily available, which diverts attention from the person and focuses it on the disability and the device. Products that employ universal design allow those that require minor accommodations the ability to use commonly available devices, thus removing social stigmas for the user, as well as increasing market share for the producer. Universal design features can often be included in a product with minimal cost, and in a subtle manner (Vanderheiden, 1990).

The lessons learned from the successes of minor product improvements due to the ADA and Telecommunications Act demonstrate that it is practical to make subtle, low-cost modifications that increase accessibility (Schneiderman, 2000; Vanderheiden, 1990). Vanderheiden and Law (2000) suggest the use of redundant methods of information presentation as an example of a simple method to increase product accessibility. The use of redundant methods to present information refers to presenting information stimuli that can be received by the visual, auditory, and tactile senses. For example, if a key on a keyboard were pressed, the tactile sensation of the key click, the sound of the key click, and the contents of the display screen updating due to the key press are redundant ways to confirm that the key had been pressed.

Newell and Gregor (1997), and Vanderheiden and Law (2000), indicate that the use of redundant methods of information presentation can assist with the use of a device in a less than optimal environment. Tactile and auditory feedback can be used in place of visual feedback when a user is unable to see the display, or when using a device in the dark. Visual and tactile feedback can be used in place of auditory feedback when the device is used in a noisy environment or by the hearing impaired. The use of redundant presentation methods is one example of how universal design can be applied (Edwards, Pitt, Brewster, and Stevens, 1995).

Through the use of simple modifications, product designers can incorporate universal design. These modifications can be subtle and inexpensive, increasing product accessibility and usability, leading to increased independence for the user, and increased market share for the producer.

2.6 - Usability

After a product is designed to be “accessible”, it must also be “usable”. Usability is defined as a composite of several usability attributes, as opposed to a single one-dimensional property. These attributes are learnability, efficiency, memorability, errors, and satisfaction (Hix and Hartson, 1993; Jordan, 1998; Nielsen, 1993; Preece, 1993). Products that incorporate these attributes should be easy to learn and remember, allow users to accomplish tasks quickly and easily, have a low error rate, and be satisfying to use (Hix and Hartson, 1993; Jordan, 1998; Nielsen, 1993; Preece, 1993).

To many, the terms usability, and universal design are synonymous. Although these terms seem similar, the concept of each is very different. “Usability” refers to the measure of “quality” and “satisfaction” related to the interaction with a device (Benyon et al., 2001). “Universal design” refers to the ability to allow many users with a wide range of abilities to access the features of a device. Products rated high in usability may allow users without disabilities the ability to accomplish tasks efficiently and with a high level of satisfaction, while those with disabilities are not able use the device at all. A device with universal design features may allow those with disabilities to access product features, but the features may be confusing, difficult to use, or not particularly satisfying to use. A usability evaluation of a product can reveal issues that in many instances can be solved with universal design features. Additionally, it is important that usability evaluations are performed on products that incorporate universal design features to ensure that these products are easy and satisfying to use.

Several methods can be used to evaluate the usability of a product. These methods include empirical studies, heuristic analysis, cognitive walk-throughs, scenarios, simplified thinking aloud and usability-expert reviews (Nielsen, 1993; Virzi, 1997).

Empirical studies are used to quantify the time and errors associated with a task. Empirical studies are generally completed in a laboratory environment, where a user is assigned a task and the time to complete the task and errors are quantified. Empirical studies often require large numbers of participants and are completed near the end of the design lifecycle (Hix and Hartson, 1993).

Heuristic analysis is a usability inspection method that requires product developers to compare the product interface to a list of usability heuristics. These heuristics, or “rules of thumb”, can assist those with little usability experience in finding and avoiding usability problems (Nielsen, 1993).

A Cognitive Walk-through is a usability inspection method that helps assess the learnability of an interface. A cognitive walk-through can help to determine if there is enough information in a user interface to learn and accomplish a goal (Virzi, 1997).

Scenarios are mock-ups of situations where users are asked to complete a specific task to achieve a specific outcome, under specified circumstances, over a certain time interval. Scenarios can be used early in the development lifecycle to elicit usability information without constructing an elaborate prototype (Virzi, 1997).

The Simplified Thinking Aloud method requires users to verbalize their thoughts while using an interface to accomplish a task. The session is transcribed and a content analysis is performed to quantify usability issues (Nielsen, 1993).

A Usability-Expert Review employs usability experts to act as the user and critique the user interface. Usability experts identify issues that would cause users difficulties by using their prior usability and system experience (Virzi, 1997).

Usability inspection methods can be employed early in the product development lifecycle, or after the product is developed (Nielsen, 1993). Evaluations that occur early in the product development lifecycle are called “formative evaluations”, while those that occur near the end of the development lifecycle, or after the product is developed are called “summative evaluations” (Hix and Hartson, 1993). Both types of evaluations use different data collection methods, and can be used to elicit different usability information (Nielsen, 1993).

Formative evaluations are usually “quick and dirty” evaluations, designed to expose the most severe usability issues early in the development lifecycle (Nielsen, 1993). By exposing these issues early in the development lifecycle, design changes can be made in time for the next iteration of the development cycle (Hix and Hartson, 1993). Due to the limited time available during product development however, formative evaluations use small pools of subjects to assess usability (Nielsen, 1993). Since formative evaluation occurs early in the development lifecycle, several evaluations and redesign cycles can be completed to ensure most usability issues are addressed prior to product release (Hix and Hartson, 1993).

Summative evaluations are generally empirical in nature, often evaluating the performance of a device, or comparing it with another (Nielsen, 1993). Qualitative methods such as questionnaires and rating scales are often used in conjunction with empirical methods to elicit as much information about usability as possible (Hix and Hartson, 1993). Since summative evaluations occur near, or sometimes even after the production phase of the product lifecycle, major usability concerns cannot be addressed until the next generation of the product is designed (Hix and Hartson, 1993). This position in the product lifecycle allows usability evaluators to complete a more formal usability evaluation, often using a large pool of participants. This results in a test that provides a more robust statistical analysis, however any recommendations cannot be implemented until the next product development cycle (Hix and Hartson, 1993).

Formative and summative evaluations are generally performed in a formal usability lab (Nardi, 1997). Usability research in this formal setting however, is not always practical. Large companies developing several products at the same time may not have adequate lab space to perform testing. Logistical constraints, such as participant recruitment and scheduling, also limit testing (Nardi, 1997).

These constraints, in addition to the desire to collect usability data in actual user environments, have lead to the use of ethnographic methods. Ethnographic methods are used to evaluate user needs and actions by observing users in their own environment (Rose, Schneiderman, and Plaisant, 1995). The behavior that a user may exhibit in their own environment may be different than what they would exhibit in a formal laboratory setting (Nardi, 1997; Rose et al., 1995).

Ethnographic methods were originally used in anthropology to study other cultures (Nardi, 1997). In an ethnographic study, anthropologists become part of the culture of interest. After a period of time, the observed culture accepts the presence of the observer, and goes about their daily lives. The observer, now an accepted part of the culture, is then able to observe behaviors that may not be exhibited in a laboratory setting (Nardi, 1997; Rose et al., 1995). Ethnographic studies in anthropology generally take a year to complete due to the observer adjusting to a foreign environment. A year-long study is not feasible for product development, however, due to the short product development lifecycles (Rose et al., 1995). If properly constructed, an ethnographic study in a domestic environment can be completed in as little as 6 weeks (Nardi, 1997).

Product manufacturers, recognizing the shortcomings of other usability evaluation methods, are employing ethnographic methods to study user needs and behaviors in their context of use. Ethnographic studies can elicit this information, which is important for creating design solutions (Nardi, 1997). Smith-Jackson, Williges, Kwahk, Durak, Capra, Nam, and Ryu (2001) found that ethnographic approaches are useful when working with older users with limited capabilities to travel outside of their communities. In particular, Smith-Jackson et al. (2001) used a remote ethnographic method that allowed participants to be involved in usability evaluations from their own homes.

Ethnographic methods can elicit a large volume of information. According to Nardi (1997), sorting through this information to find usability issues can be a great challenge. To limit the volume of data that must be analyzed, the critical incident technique can be used. This technique allows for the collection of specific usability-related data, reducing the volume of data that must be analyzed to classify usability problems.

Flanagan (1954) originally proposed the critical incident technique to collect information on observed behaviors in a variety of contexts, such as aviation, industrial operations, and computer use. This technique has been modified by several researchers, and recently adapted by Hartson and Castillo (1998), Thompson (1999), and Smith-Jackson et al. (2001), to evaluate the usability of computer interfaces in remote environments. A critical incident in the context of usability is defined as an event or interaction during task performance that indicates something either positive or negative about usability (Hartson and Castillo, 1998).

Critical incident data can be collected on a remote computer using activity recording software, such as keystroke and screen recorders that record user actions and system responses as they occur. Due to the large number of system resources required to operate this software, however, it is usually deactivated during normal computer use. To provide robust data for a critical incident report, users must activate the software prior to the start of the critical incident. According to Hartson (2001), users may have difficulty recognizing when a critical incident begins to occur. When the user realizes that a critical incident has occurred, the event has already passed and cannot be captured. Due to this, Hartson (2001) recommends that users first experience the critical incident, activate the activity recording software, and then reenact the incident by repeating the events that lead up to it.

The present notebook computer and cellular telephone studies used several of the usability methods described above. In order to determine the needs of those with visual and upper extremity disabilities, it was important that these devices were studied in their context of use. Additionally, due to limited study resources, it was important that specific usability information was collected. The notebook computer study used a remote ethnographic method in combination with a user-reported critical incident method to collect specific, context-related usability data. Since cellular telephone technology does not provide a means to collect usability data outside of the laboratory, as does the computer, participant use must be studied in the laboratory. The cellular telephone study used empirical methods, along with qualitative methods to evaluate usability. Since the study was conducted on production models of notebook computers and cellular telephones, it can be considered a summative evaluation. The data from this summative evaluation can be used to create product-specific design guidelines that can be applied to the design of future products to enhance usability.

2.7 - Design Guidelines

Product usability and accessibility can be addressed through the use of design guidelines. Design guidelines are intended to summarize human performance data, and provide suggested practices to follow when designing a user interface (Campbell, 1996). The use of guidelines can help ensure customer needs are incorporated into a product, and may shorten development time (Rosenzweig, 1996). Guidelines can be used during the design phases of a product, or to evaluate the product after design is complete (Campbell, 1996).

Several sources of design guidelines are available. Guidelines can be referenced from sources such as the popular *Guidelines for Designing User Interface Software* by Smith and Moser (1986), the *Access to Telecommunications Equipment and Customer Premises Equipment by Individuals with Disabilities* by the Telecommunications Access Advisory Committee (1997), and the *Resource Guide for Accessible Design of Consumer Electronics* by the Telecommunications Industry Association and Electronic Industries Foundation (1997), to name a few. These documents are just a sample of the many guidelines documents available, and provide numerous general guidelines intended to be used in a variety of applications. These general guidelines may seem like common sense to many designers, however to use them effectively, careful interpretation is needed to apply them to a specific application.

Although many design guidelines exist, and many seem like common sense to designers, special care must be used when applying them. General design guidelines can be vague, may occasionally conflict, and should not be applied blindly (Sanders and

McCormick, 1993). If design guidelines are applied blindly, a designer may create a usability issue in the attempt to solve another (Hix and Hartson, 1993). Other considerations, such as trade-offs, market research, educated opinions, and a deep understanding of the user population must be used in conjunction with these guidelines to ensure they are applied appropriately (Hix and Hartson, 1993; Sanders and McCormick, 1993).

The general nature of available guidelines, the requirement to interpret and tailor the guidelines to specific situations, and the possibility of conflicts within the guidelines, may lead designers to avoid using them in a product design. According to Carter (1999), designers may not use general guidelines since they often state what should be done at a high-level, without explaining how to do it for a specific application. Additionally, designers may not appreciate why a guideline exists, if they are not provided with the rationale behind it. Due to the number of general guidelines available, and the conflicts within guidelines, designers may also have trouble choosing the appropriate guideline, and deciding among alternative or conflicting guidelines (Henninger et. al, 1995). The use of product-specific guidelines, however, may reduce these concerns, resulting in much greater usage in future designs (Campbell, 1996; Phillips, 1993).

Product-specific guidelines are guidelines intended for a specific product type or family of products. Since these guidelines are tailored to the specific product type, designers are not required to spend as much time interpreting general, often abstract guidelines, choosing among alternative guidelines and dealing with potential conflicts, or deciding how they will incorporate the guidelines into the design. The use of product specific guidelines can not only save time and ease the implementation for designers, but can assist them in understanding different user classes, the task domain, and the type of tasks users perform (Henninger et. al 1995).

Guidelines tailored to a specific product are not always enough, however. According to Carter (1999), designers prefer guidelines accompanied with the rationale behind them, as well as specific examples of how the guidelines are used to assist in implementation. Without the rationale and specific examples of how to implement guidelines, designers may not be able to interpret the guidelines appropriately, and implement them effectively. Human factors and usability professionals should thus provide designers with product-specific guidelines, and include design rationale, and implementation examples whenever possible.

2.8 - Notebook Computer and Cellular Telephone Research

There is a great deal of research related to the use of notebook computers and cellular telephones. Studies have addressed user preferences for keyboards and display attributes, performance related to data entry techniques, and preferences for specific design attributes. Akagi (1992) studied user preference for key activation forces on computer keyboards, and determined that users prefer keyboards with low activation forces, although the keys with low activation forces produce more errors. A study of user performance with reduced-size numeric keypads revealed that users performed better with a full-sized keypad, and if a full size keypad was not available, performance was best when using a reduced-size keypad with wide keys (Loricchio and Lewis, 1991). Somberg (1990) studied display font attributes and determined that there was no significant difference in user performance using two similar fonts, one of which conformed to the ANSI/HFS 100-1988 standard, and another similar font that did not. Serafin, Wen, Paelke, and Green (1993) determined that there were differences in the time to dial numbers using a cellular telephone, depending on the length of the number and whether the number was familiar or not. Participants dialed unfamiliar 7-digit numbers more quickly than unfamiliar 11-digit numbers, but there was no significant difference in the time it took to enter 7-digit and 11-digit numbers that were familiar. A study on user preferences on the density of cellular telephones revealed that users prefer models with low density, and often perceive low density models as being of higher in quality than models with a higher density (Volaitis, Chou, and Wiklund, 1998).

Although there is a great deal of research involving the use of notebook computers and cellular telephones, research involving the use of these devices among those with disabilities is lacking (Newell and Gregor, 1997). In an effort to better understand the technology needs of those with disabilities, Smith-Jackson, Mooney, and Nussbaum (2001) conducted a survey-based study investigating notebook computer and cellular telephone use among those with visual and upper extremity physical disabilities. The study surveyed participants to determine specific needs and difficulties that those with disabilities face when using this technology. Ten participants with differing visual and physical disabilities were interviewed. Interviews were transcribed and the content analyzed.

The notebook computer study used two commercially available notebook computer models, both produced by prominent manufacturers. One model of computer included several subtle universal design features, while the other did not (Smith-Jackson et al., 2001).

The cellular telephone study used eight commercially available cellular telephone models, all produced by prominent manufacturers. The keypad, display, physical size, and weight of each telephone varied (Smith-Jackson et al., 2001).

In the notebook computer portion of the Smith-Jackson et al. (2001) study, only two of the ten participants had experience with notebook computers. All participants, however, were familiar with desktop personal computers, and used them on a daily basis. Participants with visual disabilities used screen readers or screen magnification tools on their home computers. Participants with physical disabilities did not use any assistive technology with their home computers.

During the study, participants were allowed to examine the computers and comment on certain features such as the screen and the keyboard layout. Participants with visual disabilities believed they would have difficulty using the computers due to low contrast between the keyboard characters and key colors, the lack of tactile identification on important keys, and the lack of a separate 10-key keypad. According to these participants, the use of tactile identification and the use of the 10-key keypad is very important. Tactile information helps users with visual disabilities to orient themselves on the keyboard, and the 10-key keypad is used to control screen reading software. Participants with upper extremity physical disabilities indicated that the lack of standardization among key locations, and compressed size of the keyboard would cause difficulty. Although participants indicated that both models of computer may be difficult to use, it was believed that the model that did not incorporate universal design features would be the most difficult, and least satisfying to use (Smith-Jackson et al., 2001).

The cellular telephone portion of the Smith-Jackson et al. (2001) study revealed that participants with disabilities use their cellular telephones primarily for safety and convenience. Participants reported that a cellular telephone is a portable communication source that allows them to contact others for help without having to search for a pay phone. Most participants used a limited number of functions, and generally made only outgoing calls. Participants with visual impairments indicated that they had trouble operating the telephones due to low display contrast, or the lack of an audio display. Most participants indicated that the telephones were difficult to hold due to the size, and that the keys were often hard to find due to size, height, and lack of tactile identification.

The Smith-Jackson et al. (2001) study revealed potential issues that users with disabilities may have when using notebook computers and cellular telephones. Since the study was limited in scope, issues were only identified and not evaluated. In order to better understand the needs of those with disabilities, it was important they have the opportunity to interact with these devices for a longer period of time. This increased interaction allowed both users, and researchers, the ability to determine and evaluate what barriers prevent those with disabilities from accessing and using this technology effectively. Once these issues are more fully identified and evaluated, simple universal design features can be included in products that allows users with disabilities to access this technology and the potential benefits this access provides.

3.0 - Research Purpose

This research explored how selected attributes of notebook computers and cellular telephones contribute to the usability of these devices among users with visual and upper extremity disabilities. Since the study was designed to evaluate the usability of specific hardware features that require physical manipulation, and not software or menu navigation, those with cognitive disabilities were not included. Although it is possible that certain cognitive disabilities may affect vision or the ability to physically manipulate objects, these participants were not included due to difficulties in recruiting and screening appropriate participants. Notebook computers and cellular telephones were selected for this study based on their prevalence and general popularity. This study was a preliminary performance study that expanded on a previous survey-based study conducted by Smith-Jackson, et al. (2001). In the Smith-Jackson et al. (2001) study, participants were surveyed about their preferences for cellular telephone and notebook computer attributes. Quantitative and qualitative data regarding the usability of specific cellular telephone and notebook computer attributes were collected in the present study. Results were used to determine what usability issues exist, and then to determine how the use of universal design principles can improve usability. These results were then used to develop product-specific guidelines that can assist user interface designers in the design of future notebook computers and cellular telephones.

4.0 - Experimental Methods - Notebook Computer Study

4.1 - Overview

In the Smith-Jackson et al. (2001) study, computer users with disabilities were interviewed to determine what their needs were when using computers. The survey required participants to have some computer experience, but notebook computer experience was not necessary.

Participants in that study indicated that it may be difficult to use a notebook computer for their computer related activities. They believed that difficulties might arise due to usability issues and conflicts between accessibility software and hardware. Participants expressed this concern during structured interviews where there was little time to fully explore the computers and determine what usability issues truly exist. Of these ten participants, only two had prior experience using a notebook computer (Smith-Jackson et al., 2001).

The concerns expressed in that survey, along with the lack of experience among the participants indicated the need to obtain more usability information. To obtain this information, it was important that users have the opportunity to interact with these devices for a longer period of time. This study was an exploratory study designed to provide those with disabilities the opportunity to further explore this technology and determine what usability problems exist.

4.2 - Design

Usability information was obtained using a remote ethnographic method in conjunction with a user-reported critical incident method. Additional data were collected through the use of subjective evaluations. Participants used two different notebook computer models, each for 30 days.

Prior to the start of the experiment, participants were asked to indicate which model of computer they preferred. A computer was then assigned to each participant for a 30-day period, in which they recorded critical incidents using activity recording software. After the 30-day period was complete, participants completed a commercially available subjective evaluation form that was customized for this study. Once the first

30-day loan period was completed, participants then received the second model of computer and repeated the process. After participants used both computers for 30-days, they were again asked which computer they preferred to determine if their preference had changed. The presentation order of each computer was balanced among participant groups to reduce biases due to learning effects, and the presentation order itself.

4.3 - Participants

Eight people were recruited from the local area to participate. These participants represented users with a wide range of abilities, including those with visual and upper extremity physical disabilities. The participant pool (Table 4.3.1) consisted of two with total blindness, two with legal blindness, two with upper extremity physical disabilities, and two with no apparent disabilities (control group). The number of participants was chosen based on the limited availability of those with disabilities in the area, and the limited availability of notebook computers for use in the study. All participants had at least a basic knowledge of computer use, and used a personal computer at home or at work.

Table 4.3.1 - Notebook Computer Study Participants

#:	Disability Group:	Age:	Gender:	Disability Description:
1	TB	41	F	Total blindness
2	LB	19	M	Legal blindness
3	LB	25	F	Legal blindness
4	ND	21	F	No disabilities
5	ND	21	M	No disabilities
6	TB	47	F	Total blindness
7	SU	20	M	Cerebral palsy, very limited use of left hand.
8	MU	26	M	Minor hand tremors when performing motor tasks

Participants with visual disabilities were included based on self-reports of legal or total blindness. Legal blindness was defined as visual acuity of 20/200 or less in the best eye with corrective lenses, or a field of view of 20 degrees or less. Total blindness was defined as the lack of any visual perception (*Employment Statistics for People Who Are Blind or Visually Impaired: 1994-1995*, 2001).

Participants with upper extremity physical disabilities were included based on self-reports of limited use of one or both hands, arms, or shoulders. Limited use in this study referred to poor manual dexterity, reduced range of motion, reduced motor control, or a reduced ability to grasp objects. It was not feasible to include those without use of their hands, since operation of these devices required that they be held and manipulated.

Users with no disabilities were included based on a self-report of no visual or upper extremity physical impairments. Since participants with disabilities in the Smith-Jackson et al. (2001) study had very limited notebook computer experience, participants with no disabilities were also included based on very limited notebook computer experience.

Although participants were included in groups of two based on disability status, the intent of this study was not to evaluate computer use solely for a particular disability category. Instead, the intent was to create a pool of participants that may have specific needs due to a disability. This pool of participants can provide overlapping usability data to identify common usability problems. Once identified, product-specific guidelines can be created to provide designers with methods in which to include universal design. These guidelines can be implemented in future product designs, reducing the usability problems, and allowing this technology to be more usable by a broad consumer market.

4.4 - Apparatus

Two different notebook computer models were used in this study, the IBM ThinkPad T-22, and the Toshiba Tecra 9000. Activity-recording software in the form of screen capture and audio recording software (Camtasia ©) was installed to record critical incident data. The software was configured so participants could provide a reenactment of critical incidents, along with detailed descriptions. Participants recorded critical incidents by pressing a specified hotkey (F11) to activate the software, reenacting the critical incident and providing a detailed description by speaking into the computer microphone, then saving it by pressing another hotkey (F12). The Questionnaire for User Interface Satisfaction (QUIS v7.0 ©), was modified to include additional sections, and used to collect subjective information at the end of the study. Participants were provided with an external storage device to transfer files between their personal computers and the notebook computers. Accessibility software, such as screen readers or screen magnification tools were installed on each machine to assist participants, as needed.

4.5 - Procedure

Participants read and signed an Informed Consent form (Appendix A) in conjunction with Virginia Tech Institutional Review Board requirements. Participants with visual impairments had the form read to them. Once the informed consent procedures were completed, participants were briefly shown each computer, familiarized with the features of each computer, and were then asked which model they preferred. Each computer shown to the participants had the brand and model names concealed to reduce selection bias due to previous experience with, or knowledge of, each manufacturer's products. After participants indicated their preference, they were assigned the first computer (determined by the experimenter prior to the experiment), an external storage device, case, and accessibility software (if necessary). Participants were asked to use the computer as they normally would, or at least five times per week, recording critical incidents as they occurred. Since this study used a remote ethnographic method, participants were allowed to use the computer however and wherever they wished.

A critical incident in this study was defined as an event or interaction that occurs during task performance that the user deems as "surprisingly positive" or "surprisingly negative". Since this study is exploratory, users were not constrained by what could be considered a critical incident, and were free to interpret a critical incident as anything they wished. This was particularly important, since different users may have different opinions of what constitutes a usability issue. Additionally, if participants were actual end users, any difficulties would be encountered after the product is purchased, and there may be a limited ability to return the product, or receive adequate technical assistance to solve them.

When a critical incident occurred, participants were asked to record it by activating the Camtasia © software and reenacting the events that lead up to it. A detailed narrative could be provided as necessary using the computer microphone and the audio recording capabilities of the software. Participants were reminded on a weekly basis to report critical incidents as they occurred.

After the first 30-day period was complete, participants completed the QUIS and then returned the computer. Participants were then assigned the second model of computer and asked to follow the same procedure, using the computer at least five times per week, and recording critical incidents as they occurred. After the second 30-day period was complete, participants again completed the QUIS, and then returned the computer and other loaned materials. Participants were then asked which model of computer they preferred, and why they preferred it.

4.6 - Analysis of Notebook Computer Study Data

The contents of the critical incident data for each computer type were analyzed using Hypertext Research quantitative analysis software, with critical incidents being categorized by computer attribute, and usability issue. Critical incident categories include, but are not limited to design issues related to the keyboard, pointing device, display, system software, and accessibility software. The number of critical incidents in each category was tallied, and then compared by computer model.

Parametric statistics were used to analyze participant preferences and subjective evaluation data. In order to analyze the subjective evaluation data from the QUIS, the mean of each section of QUIS was calculated, and a series of one-way ANOVAs were used to determine if the participant's satisfaction with aspects of one model of computer was significantly different than the same aspects of the other model of computer.

At the beginning of the experiment, participants were asked which computer they preferred. After using both computers for 30 days, participants were then asked again which computer they preferred. The preferences for each model of computer collected before and after the experiment were then compared to determine if participants changed their preference. No formal statistical tests were performed on this data, due to the small sample size and lack of changes in preference.

Analysis for the QUIS was conducted with a significance level of 0.10 ($\alpha = 0.10$) to minimize the Type I error, or likelihood of rejection of the null hypotheses when the null hypotheses are true. The null hypothesis (no difference in ratings between computers) was rejected if the probability of a chance outcome ("*p*" value) was less than or equal to the significance level (α). The significance level of 0.1 was also chosen to increase power since the sample size was relatively small, and to increase the likelihood of finding significant effects that may warrant further investigation.

5.0 - Results – Notebook Computer Study

The results of the before and after preferences, QUIS scores, and critical incidents were compiled and analyzed after participants completed the study. Due to the limited study population, the QUIS results were especially sensitive to the responses of each participant. Caution must be exercised when interpreting these results due to this sensitivity. Additionally, "Overall" measures, which are averages of all the data points for a particular computer across all groups, are intended for use as a rough means of summarizing the data. Wide variations in scores from different participant groups may greatly affect these values, which may lead to an incorrect conclusion about performance on a particular measure.

5.1 - Before and After Preferences

Participants were asked which computer they believed they would prefer after a familiarization with each computer, before the start of the initial loan period. After participants used both computers, they were again asked which computer they preferred. Six of the eight participants preferred the computer originally chosen before the start of the experiment. One participant that originally preferred the IBM computer switched their preference to the Toshiba computer. The participant indicated that a preference for the case color and design led to the change in preference. One participant indicated no preference for either computer (Table 5.1.1).

Table 5.1.1 – Preferences Before and After Loan Periods

#:	Disability:	Before Preference:	After Preference:	Change:
1	Total blindness	IBM	IBM	No
2	Legal blindness	IBM	Toshiba	Yes
3	Legal blindness	Toshiba	Toshiba	No
4	None	Toshiba	Toshiba	No
5	None	Toshiba	Toshiba	No
6	Total blindness	IBM	IBM	No
7	Physical disability	IBM	IBM	No
8	Physical disability	IBM	No preference	No

5.2 - Critical Incidents

Critical incidents were compiled by negative and positive critical incidents (Tables 5.2.1 and 5.2.2 respectively). Critical incident tables reflect issues with both computers that were related to both hardware and the user interaction design.

Although one computer may have a lower frequency of critical incidents for a given code, caution must be used in judging one computer as more “usable” than the other based on that specific code. Users may have had the same usability issues with both computers, but may have only reported it with one. Additionally, some usability issues may be strictly software related, thus inflating the number of critical incidents for a computer.

Many participants had trouble with the keyboard layouts, since they differed from that of a desktop computer. Although both keyboards differed from that of a desktop computer to some extent, the IBM keyboard differed only slightly, whereas the Toshiba keyboard was substantially different. Differences included non-standard key locations and arrangements, inconsistent key sizes, and keys that were omitted.

Neither computer provided users with descriptive feedback, or guided users to an acceptable solution to resolve problems when errors or program failures occurred. Users experienced problems with Internet connections, the operating system, applications, and peripherals. When a failure would occur, minimal feedback to the user was provided, which was often not enough to guide the user to a solution.

Several participants did not like the pointing device on either computer. Although the IBM pointing device received more critical incidents, many of the same issues may have pertained to the Toshiba computer, but were not reported by participants. Participants had difficulty controlling the pointer and accurately positioning the cursor, moving objects, etc. In addition to the movement of the pointer, some participants had trouble using the buttons on the Toshiba machine. The buttons are placed in a vertical arrangement, rather than the horizontal side by side arrangement common on most pointing devices. Additionally, there were two small buttons that controlled functions that users were unaware of, or unaccustomed to. The IBM pointing device had one large button that controlled a function users were unaware of, or unaccustomed to.

Table 5.2.1 - Negative Critical Incidents

Critical Incident Code:	IBM Frequency:	Toshiba Frequency:	IBM Issues:	Toshiba Issues:
Keyboard not what user expects	16	28	Users expect <Ctrl> key where <Fn> key is.	<Alt>, <~>, <Ins>, <Del>, <Ctrl> keys not where users expect. Sizes of <Tab>, <Caps Lock>, and <Enter> key are a different size than what user expects.
Lack of feedback to user	20	25	Trouble with internet connections, and use of DVD drive. System crashes or programs fail without providing information on how to fix.	Troubles with internet connections, and Camtasia. Several operating system crashes. Lack of battery level indicator.
Pointing device difficulties	17	11	Mouse difficult to use, difficult to see.	Mouse difficult to use. Buttons difficult to use. Unconventional button layout. Mouse difficult to learn.
Inadvertent key activation due to design	8	16	Trouble with keyboard shortcuts since users hit the <Fn> key instead of the <Ctrl> key.	Users unable to complete keyboard shortcuts due several non-standard key locations.
System state not apparent to user	10	15	Software abruptly stops working. Problems with the internet. Computer will not shut down properly.	Power saving functions cause system changes without warning user. Programs do not work properly when computer returns from power saving modes. User puts computer in power saving mode, computer reboots when user tries to return from that mode.
Does not guide user to solution	9	13	Trouble searching for files. Trouble with DVD drive / software.	Programs fail unexpectedly without providing a solution to prevent failures. Trouble accessing internet, although computer shows active connection.
User cannot accomplish task using chosen method	5	7	Internet content cannot be received in chosen manner with America Online (AOL). Unable to install preferred DVD software.	Users cannot complete keyboard shortcuts with both hands since there is only one <Ctrl> key. No floppy drive.
Case design	7	3	Trouble with case latches.	Bottom of computer gets hot, volume control hard to find.

System performance	5	3	System lag for normally fast operations. Program opening and closing is slow. System boot time is very slow.	Slow to boot. Some users unsure if computer was on due to boot speed.
Unexpected mode changes	5	3	Troubles with Internet connection (AOL). Troubles with Camtasia.	Accessibility software not active when computer returns from sleep / standby modes. Computer goes into power saving mode unexpectedly.
Display quality	4	2	Glare, viewing angle issues.	Glare, viewing angle issues.
Multimedia quality	1	3	Music is garbled, not clear.	Poor speaker sound.
Learnability	0	6		Users continued to make same errors with keyboard throughout use period.
Peripheral performance / system status	6	0	Not sure if CD drive is working correctly. Unexpected shutdowns during routine operations.	
Design for error prevention	0	5		Non-standard keyboard configuration contributes to errors when users try keyboard shortcuts.
Color contrast	3	0	Difficulty seeing pointing device.	
Lack of peripherals	0	2		Lack of floppy drive does not allow users with total blindness to fully install accessibility software.
Icon size / target detection	1	0	Trouble finding icons and pointer on screen.	

Table 5.2.2 - Positive Critical Incidents

Critical Incident Code:	IBM Frequency:	Toshiba Frequency:	IBM Issues:	Toshiba Issues:
Overall user satisfaction	10	10	Total surprise in ease of use. Machine worked smoothly. Sound output was good. Machine enjoyable to use.	Machine ran smoothly. Screen was brighter than other model. Less screen glare than other model. Opening the computer with one hand is easy.
Keyboard use	2	5	Keyboard easy to use.	<Windows> key helpful. <Ctrl> key located on far lower left was easy to find. Appreciated dots on function keys. Liked position of arrow keys.
Mouse use / preference	3	4	Scroll button is useful. Easy to use with one hand. Buttons easy to find and use by touch.	Preferred over other model. Seemed more stable and easier to get used to than other model. Liked “Back” button.
Case design	1	5	Looks neat and sleek.	Easy to open, especially with one hand. Light for active <Caps Lock> key is helpful.
Reliability / performance	5	3	Exceptional performance and reliability.	Great performance and reliability.
Display properties	2	2	Screen was clear. Sharp resolution.	Picture quality was great. Screen was clear and sharp.
Multimedia	0	3		Surprised at quality of multimedia. Sound worked really well. Sound and picture quality were great.

Power saving modes on both computers caused difficulty for many participants. When computers went into a power saving mode, such as the “Sleep” or “Standby” modes, they did not always return from those modes with a stable operating system, or all the applications running as before. In several instances, participants that used accessibility software (those with legal and total blindness) found that the software was not active when the computer returned from a power saving mode. These participants often turned the computer completely off, and restarted it to ensure that the accessibility software was reactivated.

Participants expressed dissatisfaction with the case designs. The IBM computer case has two latches that must be operated in order to open the computer, instead of the one latch on the Toshiba computer. Several participants commented on the inefficiency of operating two latches, and how opening the computer using one hand would be difficult. Participants also commented on the design of the CD drive eject button on the IBM computer, since it was easy to activate inadvertently, interrupting the user and application. One participant commented on the amount of heat generated by the Toshiba computer, while another commented on trouble finding the volume control. The participant that could not find the volume control believed this would not be a problem once they were reminded of where it was. All participants were shown the location of all the controls and peripheral ports on each computer prior to the first loan period.

Several participants noted issues with the displays on both computers. Participants commented that glare was a problem on both computers, as was the ability to view the display at different angles.

Positive critical incidents included reports of overall user satisfaction, helpful keyboard features, and system reliability. Several participants were surprised at the quality of sound, screen resolution, and the relatively smooth performance of both computers. Participants with total blindness appreciated the <Windows> key on the Toshiba computer, since it provided them with an easy way to open the “Start” menu in Microsoft Windows. These participants also appreciated tactile feedback in the form of Braille dots on the <F4> and <F8> keys.

Participants seemed to express greater satisfaction with performance and reliability of the IBM computer throughout the study. Technical issues were present with the Toshiba computer in the first half of the study, resulting in poor performance and reliability. Once these issues were resolved, the performance and reliability of both computers was similar.

5.3 - Questionnaire for User Interface Satisfaction

After participants used a computer for 30 days, they completed the questionnaire for user interface satisfaction (QUIS). These results, compared by computer and disability group, are shown below (Tables 5.3.1, 5.3.2, and 5.3.3). Text explanations are organized by QUIS section, such as “Overall Reaction”, “Screen”, etc., with explanations for the scores for each disability category explained separately.

Disability categories included Visual, Physical, None, and Overall. Each category consisted of the mean measures for participants for that specific group. Participants with total and legal blindness were included in the Visual category, those with physical disabilities were included in the Physical category, those with no apparent disability were included in the None category, and the mean of all participant scores were included in the Overall category.

The sections “Technical Manuals”, “Online Help”, and “Teleconferencing” were removed from the analyses. Most participants did not use the technical manuals, online help systems, or teleconferencing features, so a comparison between computers was not possible.

Overall Reaction

All participant groups rated the IBM computer more highly than the Toshiba computer. This difference in rating, however, was small, except for participants with legal blindness. Participants with legal blindness rated the IBM significantly higher on Overall Reaction ($p < 0.05$). When the ratings of all participant groups were collapsed into the Overall category for each computer however, the difference in ratings between the computers was not significant.

Screen

Participants with no disabilities rated the Screen section the highest on the IBM computer, while those with visual (legal blindness) and physical disabilities rated the Toshiba the highest. The overall ratings for both computers were very similar. Differences in ratings overall, and for each disability category were not significant.

Terminology

Participants with visual and physical disabilities rated the Toshiba higher in the System Terminology section, while participants with no disabilities rated both computers similarly. When participant groups were collapsed into the Overall category, the ratings for the Toshiba computer were the highest. Differences in ratings overall, and for each disability category were not significant.

Learning

Participants with physical disabilities rated the IBM the highest for Learning, while those with no disabilities rated the Toshiba the highest. Ratings for those with visual disabilities were very similar for both computers. When participant categories were collapsed into the Overall category, ratings for the IBM computer were the highest. Differences in ratings overall, and for each disability category were not significant.

System Capabilities

Participants with visual disabilities rated the Toshiba computer the highest for System Capabilities, while those with physical disabilities and no disabilities rated the IBM the highest. When ratings for each participant group were collapsed into the Overall category, the ratings for both computers were very similar. Differences in ratings overall, and for each disability category were not significant.

Multimedia

Overall ratings for Multimedia were the highest for the Toshiba computer. Participants with no disabilities rated the Toshiba the highest, while those with visual and physical disabilities rated both computers nearly the same. Participants with legal blindness rated the Toshiba slightly higher, while participants with total blindness rated the IBM higher. When participants with legal and total blindness were collapsed into the Visual category, the differences in ratings were negated, leading to very similar ratings for each computer. Differences in ratings overall, and for each disability category were not significant.

Software Installation

The IBM computer was rated the highest overall for Software Installation. Participants with total blindness and no disabilities rated the Toshiba the highest for Software Installation, while those with legal blindness and physical disabilities rated the IBM the highest. When participants with legal and total blindness were collapsed into the Visual category, the ratings for the Toshiba were the highest. Differences in ratings overall, and for each disability category were not significant.

Although participants with total blindness rated the Toshiba computer the highest in this category, it is important to note that accessibility software was installed on both computers by the experimenter before each loan period. The decision to preinstall the software for this group was made to ensure the software was installed correctly, and that participants would not quit the study prematurely due to the inability to install accessibility software. The installation of accessibility software on both computers was difficult. The IBM computer required the installed CD drive to be removed and replaced with a floppy drive to complete the installation, while the Toshiba computer required an external USB drive, and an additional driver to complete the installation. Neither the floppy drive, nor the required driver was included with the Toshiba computer. Had participants with total blindness been required to install accessibility software themselves, the ratings for this section would most likely be different.

Keyboard Use

The IBM computer received the highest ratings overall, and for each disability category for the Keyboard Use section. Differences in ratings between computers overall, and for each disability category, however, were not significant.

Pointing Device Use

The overall ratings for pointing device use were nearly the same for both computers. Participants with visual disabilities and no disabilities rated the Toshiba the highest, while participants with physical disabilities rated the IBM the highest. When combined overall, the ratings for each computer were nearly the same. Differences in ratings overall, and for each disability category were not significant.

Miscellaneous Hardware

Participants with legal blindness rated the IBM significantly higher than the Toshiba ($p < 0.07$), leading to a higher rating for the IBM by participants in the Visual disabilities category. Participants with physical disabilities also rated the IBM the highest, while those with no disabilities rated the Toshiba the highest. The IBM computer received the highest rating overall, although ratings for the IBM and Toshiba were very similar. Although legally blind participants rated the IBM significantly higher than the Toshiba, when collapsed into the Visual category, differences in ratings were not significant. Differences in ratings overall, and for the other disability categories were not significant, either.

Accessibility

Participants with visual and physical disabilities rated the IBM the highest for Accessibility, while participants with no disabilities rated the Toshiba the highest. When the disability groups were collapsed into the Overall category, Accessibility ratings for the IBM were the highest. Differences in ratings overall, and for each disability category were not significant.

Rank Sums

To determine which computer scored the best across all the measures, the mean score for each computer was rank ordered by section for each disability category (i.e. Visual, Physical, and No Disabilities), with the highest rating ranked "1", and the lowest ranked "2". These ranks were then added and summarized for each computer. The rank sums for each computer were determined by adding the rankings for each computer across all sections. The "Sum of Ranks" column contains the sum of the ranking scores for each computer. A lower sum of ranks score denotes higher ratings across all sections of the QUIS (Table 5.3.4). Although the IBM computer had the best overall rank sum score (lowest), the difference in rank scores between computers was minimal.

Table 5.3.1 - Overall QUIS Results

Model:	Overall Reaction	Screen	Terminology	Learning	System Capabilities	Multimedia	Software Installation	Keyboard Use	Pointing Device Use	Misc. Hardware	Accessibility
IBM	7.0	6.9	6.2	6.9	6.0	5.3	5.9	5.9	5.2	6.8	6.9
Toshiba	6.4	6.8	6.7	6.7	6.0	5.6	5.0	5.6	5.2	6.7	6.7

Table 5.3.2 - QUIS Results for Participants with Visual Disabilities

Model:	Overall Reaction	Screen	Terminology	Learning	System Capabilities	Multimedia	Software Installation	Keyboard Use	Pointing Device Use	Misc. Hardware	Accessibility
IBM	6.7	6.2	5.4	6.1	5.5	4.7	4.6	5.4	5.5	6.5	6.7
Toshiba	5.9	6.5	6.4	6.1	6.0	4.7	5.4	5.3	6.6	6.0	6.2

Table 5.3.3 - QUIS Results for Participants with Physical Disabilities

Model:	Overall Reaction	Screen	Terminology	Learning	System Capabilities	Multimedia	Software Installation	Keyboard Use	Pointing Device Use	Misc. Hardware	Accessibility
IBM	6.9	6.9	6.5	7.6	7.0	4.5	7.4	5.2	5.4	6.7	7.1
Toshiba	6.0	7.1	6.7	6.0	6.2	4.6	7.0	4.6	4.3	6.0	6.0

Table 5.3.4 - QUIS Results for Participants with No Disabilities

Model:	Overall Reaction	Screen	Terminology	Learning	System Capabilities	Multimedia	Software Installation	Keyboard Use	Pointing Device Use	Misc. Hardware	Accessibility
IBM	7.8	7.5	7.5	8.0	6.1	6.8	5.3	7.4	4.6	7.6	7.2
Toshiba	7.7	6.9	7.5	8.7	6.0	8.6	5.9	7.0	4.8	8.6	8.4

Table 5.3.5 - Sum of Ranks for QUIS Results

Model:	Overall Reaction	Screen	Terminology	Learning	System Capabilities	Multimedia	Software Installation	Keyboard Use	Pointing Device Use	Misc. Hardware	Accessibility	Sum Of Ranks
IBM	3	5	5.5	4.5	4	5.5	5	3	5	4	4	48.5
Toshiba	6	4	3.5	4.5	5	3.5	4	6	4	5	5	50.5

6.0 - Discussion – Notebook Computer Study

Previous research by Smith-Jackson et al (2001) revealed the needs and preferences for the use of notebook computers among those with visual and upper extremity physical disabilities. Although participants in that study had experience with personal computers, very few had experience with notebook computers. This study expanded on that research by providing participants experience with notebook computers within their context of use.

A remote-ethnographic method and critical incident reporting technique was employed to collect usability data. Participants were asked to use each computer for 30 days and record any usability issues that occurred using the activity recording software that was provided (Camtasia ©). The Questionnaire for User Interface Satisfaction (QUIS v.7.0) was then used to collect supplemental data, and participants were asked about their computer preferences before and after the study.

6.1 - Interpretation of Results

In the Smith-Jackson et. al (2001) study, participants indicated that non-standard key locations, lack of tactile feedback on specific keys, and the lack of contrast between the keyboard and printed key characters would create difficulties in using the computers. This study confirmed that participants, especially those with visual disabilities, had difficulty using both computers due to non-standard key locations. Issues from the previous Smith-Jackson et. al (2001) study, such as the keyboard and key character contrast, and the lack of tactile feedback on specific keys was addressed through the use of different computers in this study. The computers used in this study were newly released models with improved key and character color contrast, and improved use of tactile landmarks to identify keys.

Critical Incidents

Participants reported difficulties with both keyboard layouts, since they differed from that of a desktop computer. The IBM keyboard differed from a desktop computer keyboard only slightly, whereas the Toshiba keyboard was substantially different. Consistency between the notebook computer keyboards and desktop computer keyboards is important since users switched back and forth between the two frequently (Norman, 1988; Hix and Hartson, 1993; Schneiderman, 1993; Wickens, 2000).

The IBM keyboard layout differed from that of a desktop computer keyboard in the location of the left control key (<Ctrl>). The left <Ctrl> key is traditionally located in the lower left hand corner of the keyboard. On the IBM computer however, a special function key (<Fn>) occupied this location (Figure 6.1.1). The left <Ctrl> key was located one key to the right of this <Fn> key. Users had trouble completing Windows-related <Ctrl> key shortcuts (i.e. <Ctrl-S> to save, <Ctrl-C> to copy, etc.) due to this positioning.

Figure 6.1.1 - Position of <Fn> and <Ctrl> Keys on IBM Keyboard

Although the placement of the <Fn> key caused many errors when performing common shortcut key tasks in Windows applications, special functions used in conjunction with the <Fn> key were assigned to keys that were not common Windows shortcut keys, reducing the possibility of catastrophic errors (Norman, 1988). Additionally, keys that were larger in size than the character keys on a desktop computer keyboard (i.e. <Tab>, <Caps Lock>, etc.) were larger in size than the character keys on the IBM computer. The consistency in key size to that of a desktop computer keyboard helped the users to distinguish these keys from the character keys (Sanders and McCormick, 1993). Keys that were arranged by function, such as the <Insert>, <Delete>, <Home>, <End>, <Page Up>, and <Page Down> (referred to as the "six pack") on a desktop computer keyboard, retained the original arrangement and positioning on the IBM keyboard (Figure 6.1.2). The consistency in placement for these keys was important for users with visual disabilities, since these keys remained in a position where they were expected (Sanders and McCormick, 1993; Wickens, 2000).

Figure 6.1.2 - “Six pack” Configuration on IBM Keyboard

The Toshiba keyboard was substantially different than that of a desktop computer keyboard. Several keys, such as the <Ins>, <Alt>, <~>, <Del> were in non-standard locations, away from common keyboard landmarks such as the <Spacebar>, <Tab>, or the “six pack” (Figures 6.1.3 and 6.1.4). There was only one <Ctrl> key on this keyboard, located on the left side. The right <Ctrl> key common on desktop computer keyboards was omitted (Figure 6.1.4). Keys such as the <Tab> and <Caps Lock> key were similar in size to the character keys, reducing the ability to differentiate them from the character keys by touch (Figure 6.1.5).

Figure 6.1.3 - Non-standard Locations of <Alt> and <~> Keys on the Toshiba Keyboard
(<Spacebar> is to the right of the <~> key.)

Figure 6.1.4 - Non-standard Locations of <Ins>, <Del>, <Alt>, <Home>, <PgUp>, <PgDn>, and <End> Keys on the Toshiba Keyboard. (<Spacebar> is to the left of the <Ins> key. Note lack of right <Ctrl> key.)

Figure 6.1.5 - Reduced Size <Tab> and <Caps Lock> Keys on the Toshiba Keyboard (Note <Tab> and <Caps Lock> keys are similar in size to character keys.)

Participants indicated trouble completing keyboard shortcuts in a manner they were accustomed to when using a desktop computer. Completing <Ctrl-key> shortcuts (i.e. <Ctrl-S> for save, <Ctrl-C> for copy, etc.) was difficult with both computers due to the non-standard keyboard layouts. The Toshiba computer had only one <Ctrl> key, so users with limited use of the left hand, or that wanted to perform <Ctrl-key> shortcuts using their right hand found this arrangement troublesome. <Ctrl-key> shortcuts are important since experienced users use them to perform common tasks more efficiently (Nielsen, 1993; Schneiderman, 1998). Additionally, users with disabilities (i.e. total blindness) may only know how to accomplish certain tasks using <Ctrl-key> shortcuts.

<Alt-key> shortcuts (i.e. <Alt-F> to open a file menu, <Alt-F4> to close a program, etc.) were difficult for participants, especially those with total blindness, when using the Toshiba computer. The <Alt> keys are traditionally placed in locations adjacent to each end of the <Spacebar>, which users may use as a keyboard landmark to find the <Alt> keys. On the Toshiba computer however, the <Alt> keys are located away from the <Spacebar>, and the positioning of the <Alt> keys in reference to the <Spacebar> was not consistent on each side. The left <Alt> key was located two keys to the left of the <Spacebar> (Figure 6.1.3), and the right <Alt> key was located three keys to the right of the <Spacebar> (Figure 6.1.4). Users with total blindness had trouble remembering this arrangement, and committed several errors when trying to accomplish tasks such as opening a file (i.e. <Alt-FO>), or closing a program window (<Alt-F4>). The use of a consistent key arrangement, and arranging these keys around a common landmark such as the <Spacebar>, could have reduced these errors (Norman, 1998).

The position of the <Ins> key on the Toshiba computer may contribute to catastrophic errors for participants with total blindness. The <Ins> key is located next to the <Spacebar>, where the <Alt> key is traditionally located (Figure 6.1.4). When these participants try to close applications using the <Alt-F4> shortcut, they may inadvertently press <Ins-F4>, which is the command to close the accessibility application JAWS for Windows (a screen reading application commonly used by those with total blindness). Although the occurrence of catastrophic errors due to this shortcut key combination was not clear from the critical incident reports, one participant mentioned it in conversation. Keyboards should be designed to prevent catastrophic errors when possible (Norman, 1988).

When a problem with hardware, or the user interface occurred, the feedback to users was not descriptive enough to guide users to an acceptable solution. Users experienced problems with Internet connections, the operating system, applications, and peripherals. According to Schneiderman (1993), user interfaces should be designed to provide users with feedback describing why a failure occurred, and what can be done to resolve it, rather than providing a reference to an obscure hexadecimal memory address and a suggestion to try again or call technical support. Feedback should also be provided for changes to hardware modes, preferably before the computer changes modes. One participant had several difficulties with a sudden reduction in screen brightness on the Toshiba computer. The screen automatically dimmed to preserve power when the battery was low. This was not apparent to the user, who spent time trying to restore the screen brightness, rather than taking precautions to prevent data loss due to an unexpected computer shutdown. Higher quality feedback indicating the computer was going into a power saving mode could have reduced these frustrations.

The pointing device on both machines was problematic for most participants, with the IBM pointing device receiving the majority of the critical incident reports. The Toshiba computer uses a pointing device very similar to the IBM computer, however it is unclear whether participants did not experience as many critical incidents, or just failed to report them.

Since the pointing devices on both computers were very similar, the software used to control the pointing device may have had an effect on the usability and critical incident reports. Both computers used different drivers for the pointing devices, and both had different default motion speeds. The default motion setting for the pointer on the IBM computer was set halfway between “Slow” and “Fast”, while the sensitivity setting (under the “Track Point” tab in the “Mouse Properties” window) was set halfway between “Firm Touch” and “Light Touch”. The Toshiba computer only had a control for cursor speed and acceleration, and not for sensitivity. It is unclear whether users customized the settings on each machine to suit their preferences, or if they used the default settings. Although the Toshiba computer did not receive as many critical incident reports for the pointing device, it is unclear whether users had less trouble with it, or neglected to report critical incidents. Critical incidents relating to pointing device use may have been reduced had participants adjusted these properties to suit their needs. According to Nielsen (1993), novice users may not know that these properties can be adjusted, and thus did not customize the device for optimal performance.

Participants recorded difficulties using the pointing device buttons on the Toshiba computer. The Toshiba computer uses a button arrangement that is not typical of the button arrangements on most pointing devices (Figure 6.1.6). This pointing device has two small buttons arranged horizontally, and two large buttons arranged vertically. The large buttons are used for left and right click functions, while the small buttons have special functions. The small left button serves the same function as the “Back” button in an Internet browser, while the small right button allows users to scroll up or down in an application. Since the function of these buttons can be reassigned, none of the buttons were labeled and users were unsure of their functions. There may have been some confusion in the mapping of the function of each button due to the horizontal layout of the top buttons, and since the left and right click buttons on a standard mouse are traditionally in a horizontal layout (Norman, 1988). The “Mouse Properties” application in the Windows Control panel explains the function of each button on the Toshiba computer, however users may not be aware of this application, or did not open this application to discover those functions. Nielsen (1993) reports that novice users may not explore and customize these features, which may further explain why users had difficulty using these buttons.

IBM Pointing Device Buttons

Toshiba Pointing Device Buttons

Figure 6.1.6 – Pointing Device Buttons on IBM and Toshiba Computers

Power saving modes on both computers caused difficulties for many participants. Software that was working normally before the computer went into a power saving mode occasionally became unstable when the computer returned from that mode. Participants using accessibility software such as “JAWS” and “ZoomText” (both for Windows) experienced the most difficulty, since they relied on this software to use the computer.   In many instances participants tried turning the computer off to restart it. Due to hardware designed to compliment the Windows shut down process, however, the <Power> button did not work as expected. The <Power> button on Windows-based computers is no longer a momentary contact switch, but requires users to press and hold it for several seconds to shut the computer down. This design may be intended to prevent users from shutting down the computer without using the Windows shutdown procedure. This new mode of operation, however, may not match the mental model users have of a <Power> switch, which lead to further difficulty (Norman, 1988).

Participants recorded critical incidents related to the case design on both computers. Case design issues were relatively minor for the Toshiba computer, with negative critical incidents attributed to case temperature and the location of the volume control. Several critical incidents were attributed to the design of the IBM case, however. Participants disliked operating the two latches required to open the IBM computer. The Toshiba computer has one latch that can be easily operated with one hand. Participants commented on the inefficiency of operating the two latches on the IBM computer, and how opening the computer with one hand would be difficult. Participants also expressed frustration for the design of the eject button on the CD drive on the IBM computer, since the button protruded at a slight angle, and was easy to inadvertently activate. Inadvertent activation of the CD drive interrupts the user from their current task, and may interrupt the system. When this occurs a user must stop what they are doing and close the drive door. If the application used when this occurs is being run from the CD, or if the drive is closed and the drive contains a CD configured with an "autorun" script, the user must stop and attend to the interrupted application, or to an application attempting to load. The button on the Toshiba CD drive did not protrude at an angle, which may have helped to prevent inadvertent activation.

Questionnaire for User Interface Satisfaction

Results for the QUIS differed very little between computers. Participants recorded critical incidents, and made comments about usability issues in passing conversation. When completing the QUIS, however, strong preferences for the attributes of one computer over the other were not apparent.

The lack of significant differences between the computers in the QUIS sections may be attributed to several causes. Lack of computer use, the lack of appropriate benchmark tasks to compare computers, or QUIS questions that may have been too general may have reduced the ability to compare QUIS measures.

The study was designed under the assumption that participants would be motivated to participate in the experiment, and would make a reasonable effort to transition from the desktop computer to the notebook computers. Once a technical difficulty occurred, however, notebook computer use was quickly curtailed and participants returned to using their desktop computer. When this occurred, participants may have lost motivation and did not use both computers in a similar fashion.

The lack of appropriate benchmark tasks may have contributed to the lack of differences in the QUIS scores. Benchmark tasks are used to replicate specific tasks within an interface, allowing for two or more interfaces to be compared (Hix and Hartson, 1993). The elimination of benchmark tasks was justified due to the assumption that participants would have different levels of experience, use different applications, and would use the same applications on both machines. Additionally, due to the different levels of participant experience, and the remote ethnographic nature of the study, it was very difficult to determine which benchmark tasks would be truly appropriate to compare the computers, and could be effectively administered while preserving the remote ethnographic nature of the study.

Participants mentioned many usability issues in critical incidents, and passing conversation. The questions in the QUIS, however, were not written in a manner that would capture these same issues expressed in those comments. The language used for the questions may not have been interpreted in a way that the users understood, or the questions did not ask about items related to issues the user had. For example, the "Pointing Device" section asked a question about the force required to activate the pointing device. Participants did not see this as a question asking to evaluate pointing device sensitivity, although many participants commented on pointing device sensitivity throughout the experiment. The "Keyboard" section, asked many questions about keyboard design, such as key height, key width, tactile force, etc., but did not ask questions about keyboard layout. Users did not indicate any strong opinions on key attributes such as height, width and tactile force, in critical incidents or conversation, however they did indicate strong opinions on keyboard layout. Questions about keyboard layout were not asked in the QUIS, since it would have been difficult to write objective questions that were not biased towards one model of computer over the other.

Before and After Preferences

Participant preference stayed relatively consistent after using the computers, with one participant changing their computer preference, and another participant indicating no preference. Participants with no disabilities frequently commented on the attractive case design of the Toshiba machine, while participants with visual disabilities commented on the keyboard layout of the IBM.

6.2 - Universal Design Principles

Universal design is the design of products and environments to be usable by the greatest pool of people as is practical, without the need for adaptations or specialized design (Center for Universal Design, 1997). The inclusion of universal design in products can increase market share, as well as increase compliance with regulations such as the Americans with Disabilities Act, Section 255 of the Telecommunications Act, and Section 508 of the Rehabilitation Act (Mondak 2000; Mueller, 1995).

The most common critical incidents recorded for both computers were compared with the universal design principles (Table 2.5.1) to determine how universal design can be applied to each computer to improve accessibility. Each principle is listed below, along with explanations of how each computer model demonstrated, or violated these principles.

Equitable Use

The principle of equitable use refers to a design that is useful and marketable to people with diverse abilities. A design for equitable use allows for consistent operation for all users, avoids segregating or stigmatizing users, and makes design appealing for all (Center for Universal Design, 1997).

The design of each computer model is marketable to people with diverse abilities to some extent, however both computers tend to exclude people with specific disabilities. The design of both computers was relatively appealing to all groups, however the keyboard design of the Toshiba was not appealing to those with total blindness, and the case latch design of the IBM was not appealing to any participant group.

The design of the Toshiba keyboard does not provide equitable use for skilled touch typists, users with total blindness, or users with limited use of one hand. The use of a non-standard keyboard layout hinders skilled touch typists, since these users must learn new locations for commonly used keys, or frequently stop to look at the keys while typing. Users with total blindness may not be able to confidently complete <Alt> key shortcuts, since placement of the keys is not consistent with a desktop computer, and they are placed in inconsistent locations on each side of the keyboard. Additionally, users with total blindness may hit keys inadvertently due to the new locations, which may close accessibility software, or require them to find and correct errors that may not have occurred for other users who were able to see the new key locations. Users with limited use of one hand may not be able to accomplish keyboard shortcuts like other users, since there is only one <Ctrl> key. An example of the keyboard layout used is shown below (Figure 6.2.1).

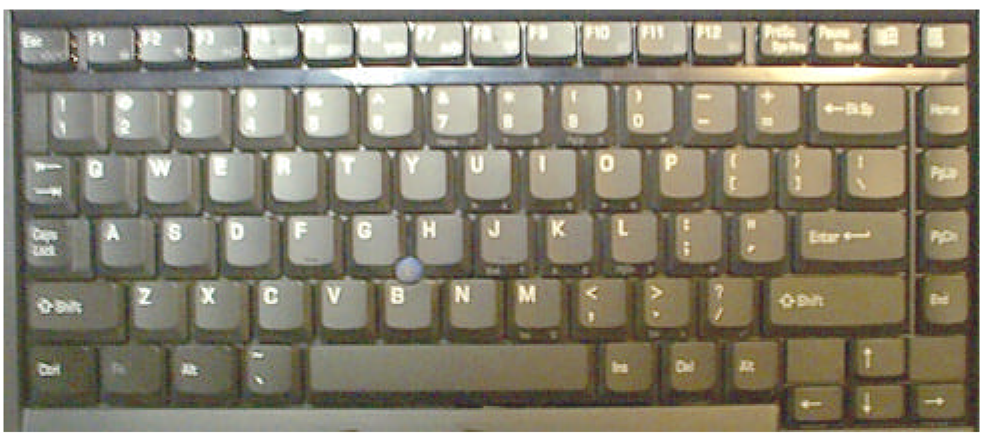


Figure 6.2.1 - Toshiba Tecra 9000 Keyboard

The design of the IBM keyboard is such that people with a variety of ability levels can accomplish work with it in a manner they may be accustomed to with a desktop computer. The IBM keyboard provides relatively equitable use for all, except for initial use by skilled touch typists or those with total blindness that rely on the use of <Ctrl> key shortcuts. The placement of the <Fn> key where the <Ctrl> key is traditionally located on the desktop computer reduces the equitable use of this design, since users must remember that the location traditionally reserved for the <Ctrl> key now contains the <Fn> key. While this may not cause catastrophic errors for these users, the use of <Ctrl-key> shortcuts is not consistent with that of a desktop computer keyboard. Since the position and reach required are similar to that of a standard desktop computer keyboard, learning the new position of the <Ctrl> key should not be too difficult. An example of the keyboard layout used is shown below (Figure 6.2.2).

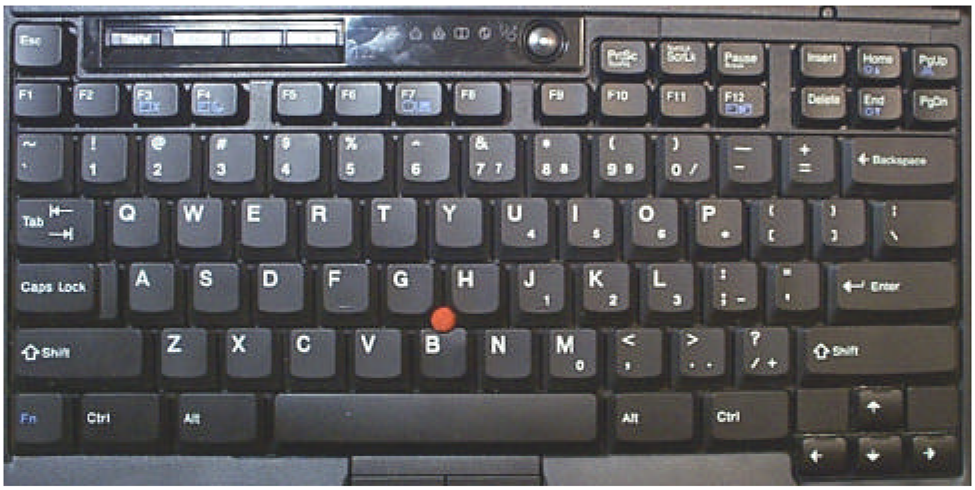
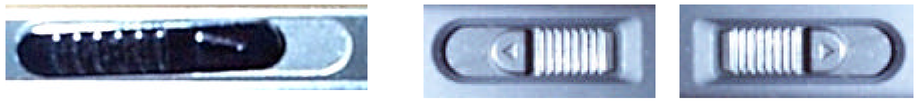


Figure 6.2.2 - IBM Thinkpad T-22 Keyboard

The design of the IBM case latches may not provide for equitable use. People with limited use of both hands may be hindered, since the operation of two latches is required to open the case. The design of the latches does accommodate for this difficulty somewhat, since simultaneous operation of both latches is not required. The latches have a mechanism that allows them to be operated sequentially, so those with limited use of both hands can still open the computer by operating one latch at a time. The latches are also recessed, which may reduce the ability to feel them for users with total blindness, or users with limited motor skills or tactile sensation (Figure 6.2.3).

The design of the Toshiba case provides equitable use for all users. Opening the case only requires the use of one latch, facilitating operation with one hand. The latch is also prominent, and easy to identify by touch, allowing users with low vision, poor tactile sensation, and poor motor skills to find and operate it easily (Figure 6.2.3).



Case latch for Toshiba Case latches for IBM
Figure 6.2.3 - Examples of Computer Case Latches

Flexibility In Use

Flexibility in use refers to a design that accommodates a wide range of individual preferences and abilities. Flexibility in use allows users a choice in the method they wish to use a device, accommodates use with either hand, or facilitates accuracy and precision (Center for Universal Design, 1997).

The IBM computer keyboard was the most flexible in use, since the relatively standard keyboard layout did not require extensive relearning, and allowed users to accomplish keyboard shortcuts using alternative methods. The IBM case latches were less flexible in use, but could be operated by users with limited use of both hands if operated sequentially with one hand.

The omission of the right <Ctrl> key on the Toshiba computer limits the flexibility in use for all users, since users must press the <Ctrl> key with their left hand, or use an awkward posture to press it with their right hand. Users with limited use of their left hand may have difficulty accomplishing <Ctrl-key> shortcuts due to this design. The case latch provided the greatest flexibility in use, since it was easy to feel, and could be operated quickly and easily with either hand.

Simple and Intuitive Use

The principle of simple and intuitive use refers to a design that is easy to use and understand, regardless of experience, knowledge, language, etc. Designs that incorporate simple and intuitive use eliminate unnecessary complexity, are consistent with user expectations, and provide effective prompting and feedback (Center for Universal Design, 1997).

According to Mueller (1995), simple and intuitive use is especially important for casual users. Users in this study did not use the computers as their primary computing platform, and only used the computer intermittently. Since the computer was only used intermittently, simple and intuitive use is important if users are to accomplish work with the system quickly and efficiently.

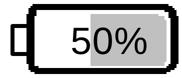
Both computers were relatively simple and intuitive to use, except for some minor hardware issues on the Toshiba computer. Most issues regarding the principle of Simple and Intuitive Use on both computers were related to the operating system.

Other than the non-standard layout on the Toshiba computer mentioned earlier, the battery level indicator was not simple and intuitive to use. Several critical incidents were reported regarding the lack of a battery level indicator. An icon to indicate battery level is present, and located in the system tray portion of the Windows task bar (lower right corner of the screen). The function of the icon, however, is not intuitive. This icon, in the shape of a light bulb, simulates a light bulb that dims as the battery level decreases through the use of different color gradations. The design of this icon, and the use of a light bulb metaphor, creates unnecessary complexity, is not consistent with user expectations, and does not provide effective feedback to the user.

The use of the light bulb icon, and the light bulb metaphor to indicate the battery level is cognitively complex, and inconsistent with a users mental model of a battery. Users must have experience using DC powered light sources, such as a flashlight, or experienced leaving their car headlights on, to understand this metaphor. Additionally, the different color gradations (light bulb intensities) used to indicate battery level may be difficult to perceive, especially for those with visual difficulties. The use of a light bulb metaphor may be an appropriate battery level indicator a portable flashlight, however, it is not appropriate for a portable computer.

The principle of Pictorial Realism dictates that such features as battery level should be communicated using an icon that simulates real world perceptions of batteries (Roscoe, Corl, and Jensen, 1981). Due to this, an icon in the shape of a small "AA" size battery should be used to indicate battery level. As the battery level depletes, the icon should indicate the estimated battery life remaining by using a status bar and label inside the icon (Figure 6.2.4). Icons such as this are commonly used on battery-powered consumer products, do not require users to interpret a complex metaphor, and may be more consistent with the user's mental model of a battery.

Icon used on Toshiba computer



Icon suggested for use on Toshiba computer

Figure 6.2.4 - Illustrations of Battery Level Icons

The Windows operating system used on both computers was relatively simple and intuitive to use, unless an error occurred. Once an error such as a program failure, or software conflict occurred, a dialog box containing a vague description of the problem and a hexadecimal memory address would appear, instructing users to retry the program, or contact technical support. The lack of constructive feedback for solving the problem did not provide for simple and intuitive use, since simply retrying the operation often resulted in the same error, and users may give up or be reluctant to contact technical support.

Perceptible Information

The principle of Perceptible Information requires that necessary information is communicated to the user, regardless of ambient conditions or sensory abilities. Perceptible information uses redundant modes of information presentation, adequate information contrast to maximize the legibility of essential information, and compatibility with techniques used by people with sensory limitations (Center for Universal Design, 1997).

The battery level icon on the Toshiba computer violated the principle of Perceptible Information. The icon is a light bulb, and uses a light bulb metaphor where different color gradations (to simulate brightness intensities) indicate battery level (Figure 6.2.4). The icon, located in the Windows system tray in the lower right corner of the screen, is relatively small, and slight variations in color are difficult to discern. An attempt is made to provide a redundant method to indicate battery level, however, through the use of a rollover text label. For this redundant presentation method to work effectively, however, users must first be aware that this icon denotes battery level, and be able to understand the wording used in the rollover text.

When a user positions the cursor over the icon, a small text box (rollover text) appears with an estimate of the power remaining. The messages that appear in the text box are not very clear or consistent. An example of this is when the battery has 34% of its capacity remaining, and the AC power pack is connected, the message reads “Full power remaining 34% On AC power”. When the power pack is not connected and the battery has 34% of its capacity remaining, the message reads “Normal remaining: 34% 1 hour 3 minutes”. The term “Normal remaining” may be unclear to users, since there is no reference to battery power.

The battery level icon should be replaced with an icon that is more consistent with a user’s mental model of a battery, and does not rely on slight color gradations to indicate remaining power level. The use of an icon like this, along with a status bar and label indicating the percent of battery life remaining, would allow users to quickly determine battery level (Figure 6.2.4). Rollover text messages should also be clear and consistent, and refer to what the icon represents. Example rollover messages, when used with the battery icon described above could be “Power remaining: 34% (1 hour 3 minutes)” when the power pack is not connected, and “AC Power. (Recharging: 34% Capacity)” when the power pack is connected.

Tolerance for Error

Designs that incorporate tolerance for error minimize hazards and adverse consequences of accidental actions. These designs provide fail-safe features, warn of errors, and discourage unconscious actions (Center for Universal Design, 1997).

The IBM computer provided the greatest tolerance for error, except for the design of the case latches, the location of the <Fn> key, and the design of button to open the CD drive. The design of the latches on the IBM case may be difficult to feel, and difficult to operate with one hand. Although the latches allow for sequential operation, if a user bumps the case inadvertently, the latches may reset and the user must operate them again in order to open the computer.

The placement of the <Fn> key where the <Ctrl> key is traditionally located on the IBM can create errors due to failed <Ctrl-key> shortcut attempts. This design does provide for some error tolerance, however, since special function keys that are used are programmed to keys that are not commonly used as <Ctrl-key> shortcuts, thus minimizing the potential for catastrophic errors. Additionally, when keys that do not have special functions are pressed in conjunction with the <Fn> key, they do not insert characters, saving the user from the need to correct additional errors inserted into their work.

The design of the button used to open the CD drive on the IBM computer does not provide for error tolerance. Although the design of this button is easy to operate, it is also easy to operate inadvertently. When inadvertently activated the drive may open, interrupting the user, and potentially interrupting the application.

The design of the Toshiba keyboard provides comparatively less error tolerance, especially for skilled touch typists and those with total blindness. Due to the non-standard key placement, touch typists that are skilled with the use of a traditional QWERTY keyboard may make errors when trying to use the <Alt>, <~>, <Ins>, <Del>, <Home>, <PgUp>, <PgDn>, and <End> keys. Users with total blindness may commit many of these same errors, however, with greater consequence. Users with total blindness that use <Alt> key shortcuts, such as <Alt-F-S> to save a file, may inadvertently press the <~> or <Ins> keys instead, thus inserting characters into their work that need to be removed, in addition to failing to accomplish the intended task. If a user with total blindness tries to close a file using the <Alt-F4> shortcut, they may inadvertently press <Ins-F4>, which closes the accessibility application JAWS for Windows. These errors may be catastrophic for users with total blindness, since they may lose a considerable amount of work when they occur.

Low Physical Effort

Designs that incorporate the principle of low physical effort can be used efficiently, and comfortably with a minimum of fatigue. These designs allow users to maintain neutral postures, use reasonable operating forces, and minimize repetitive actions and physical effort (Center for Universal Design, 1997).

The latches on the IBM case violate the principle of Low Physical Effort. The operation of two latches on this computer requires more physical effort than the operation of one latch on the Toshiba computer. The small, recessed design of these latches may require extra physical effort for those with total blindness to find, and for those with limited tactile sensation and dexterity to operate. The latch design of the Toshiba computer demonstrates the principle of Low Physical Effort. This latch is large, prominent, easy to find, and can be easily operated with one hand.

The omission of the right <Ctrl> key on the Toshiba computer violates the principle of Low Physical Effort. Users with limited use of the left hand must use an awkward posture to complete common <Ctrl-key> shortcuts located on the left side of the keyboard (i.e. <Ctrl-S> for save, <Ctrl-C> for copy, <Ctrl-V> for paste, etc.), and may not be able to complete <Ctrl-key> shortcuts located on the right side of the keyboard (i.e. <Ctrl-P> for print, <Ctrl-O> for open, etc.). The addition of a right <Ctrl> key would allow these users to complete these shortcuts without using an awkward posture.

The pointing devices on both computers violated the principle of Low Physical Effort. These types of pointing devices are generally very sensitive, and fine motor skills and coordination is needed to use them effectively. If a user has limited motor skills, a great deal of effort is required to position the cursor accurately, and to avoid overshooting intended targets.

Size and Space for Approach and Use

Designs that incorporate the principle of size and space for approach and use provide for approach, reach, manipulation, and use regardless of a user's size, posture, or mobility. Additionally, these designs provide a clear line of sight to important elements for seated or standing users, make reach comfortable for users, and accommodate different hand or grip sizes (Center for Universal Design, 1997).
Both of the computers used in this study provided reasonable size and space for approach and use by different users.

The inclusion of the universal design features outlined above should help notebook computer manufacturers produce products that are more usable, and accessible for those with disabilities, as well as those without. Products that include these features may enjoy a greater market share and allow employers to provide a consistent set of computing tools to employees regardless of disability. Additionally, the adoption of the universal design features outlined above can help product manufacturers, as well as employers comply with disability-related regulations, such as the Americans with Disabilities Act, Section 508 of the Rehabilitation Act, and Section 255 of the Telecommunications Act (Mondak 2000; Mueller, 1995).

6.3 - Usability Principles

Usability principles include learnability, memorability, efficiency, errors, and satisfaction (Hix and Hartson, 1993; Jordan, 1998; Nielsen, 1993; Preece, 1993). The most common critical incidents reported in this study were analyzed to determine which usability principles were violated. Product-specific design guidelines based on these issues, and the issues identified in the universal design section (section 6.2) will be addressed in section 6.4.

Learnability

Learnability is a measure of how quickly a user can become proficient at using the system (Nielsen, 1993). Since participants in this study were experienced computer users, the overall learnability of these systems was not a major issue.

The learnability of the non-standard keyboard layout on the Toshiba computer, however, was an issue for some participants. The non-standard placement of certain keys on the Toshiba computer created a challenge for participants to learn. This difficulty in learning was compounded when participants transitioned to and from their desktop computer and the Toshiba computer. Since the users in this study were casual users, and the notebook computer was not their primary computing platform, learnability was especially important (Mueller, 1995; Nielsen 1993).

The learnability of the IBM computer was relatively high. Since the keyboard on this computer was nearly identical to that of the desktop computer users may have been accustomed to, they did not have to learn a new keyboard layout, and could become proficient quickly.

Memorability

According to Nielsen (1993), memorability is a measure of how well a user can learn an interface, leave that interface for a period of time, and then return to complete work with it. The memorability of an interface is very important for systems that are used intermittently, or that are used by casual or non-captive user classes such as the users in this study (Hix and Hartson, 1993; Nielsen, 1993).

The Toshiba computer keyboard had low memorability due to the non-standard layout. Users had difficulty transitioning between the desktop computer keyboard, and the Toshiba keyboard due to this layout. As participants continued to use the Toshiba keyboard, errors due to this layout may have lessened, but were not eliminated. Participants that committed errors with this keyboard early in the study continued to commit the same errors later in the study. This was especially troublesome for those with total blindness, since these participants relied on procedural memory for typing, and could not look at the keys to determine their functions (Matlin, 1998). Inexperienced users that do not use other computers may learn this non-standard layout reasonably well, however it was troublesome for experienced users that use more than one computer.

Efficiency

According to Nielsen (1993), efficiency is the ability of an experienced user to accomplish tasks quickly and proficiently after learning an interface. Experienced users often use keyboard shortcuts to increase efficiency, since tasks can be completed more quickly with the keyboard shortcuts than with the mouse (Nielsen, 1993; Schneiderman, 1998). Users with total blindness, however, rely on the keyboard shortcuts to complete common tasks, regardless of gains in efficiency.

The keyboard layouts on both computers reduced efficiency. As mentioned earlier, the IBM keyboard placed a <Fn> key where the <Ctrl> key is traditionally located, requiring users to remember this new position, or retry a keyboard shortcut after a failed attempt. Fortunately, the <Fn> key does not have a function unless used in conjunction with a key in which a function is assigned. When the <Fn> key, mistaken for the <Ctrl> key, was used in combination with another key on the IBM, the result was not the insertion of an undesired character, but only a failed shortcut attempt. The non-standard key locations on the Toshiba computer, however, not only required users to retry failed shortcuts, but required users to find and correct errors inserted into their work. The requirement to correct errors due to failed shortcut key attempts reduces user efficiency.

The isometric pointing devices on both computers were relatively inefficient. Users that are accustomed to using a mouse on a desktop computer must learn how to use this new pointing device to accomplish tasks. Many users find the control of these devices difficult, and often overshoot the intended targets, fail to drag items, etc., thus reducing the efficiency of completing tasks. This finding is consistent with research by Sommerich (2002), in which users were less efficient and committed more errors when using the isometric pointing device, when compared to an external mouse. Fortunately, both computers have external mouse ports so users can avoid using this device by using an external mouse and accomplish work more efficiently.

Errors

The design of the IBM keyboard provided much greater error tolerance than the Toshiba keyboard. The design of the case on the Toshiba, however, provided greater error tolerance than the case on the IBM. A discussion of error tolerance for both computers is provided in the "Tolerance for Error" section under Universal Design Principles (Section 6.2).

Satisfaction

The results of each QUIS section can be used as a measure of satisfaction. Satisfaction ratings varied depending on the specific aspect of the interface, and the disability group. Although ratings varied by section and disability group, differences in satisfaction ratings were generally not significant.

According to Nielsen (1993), satisfaction can also be measured by asking users which system they preferred after they used it. The results of the "Before and After Preferences" portion of the study did not yield any significant difference in satisfaction between computers. Nearly all participants that preferred one model of computer before the study began, preferred the same model at the end of the study. Since there were no significant changes in before and after preferences, and the "Overall Satisfaction" ratings were relatively high (7.0 for the IBM and 6.6 for the Toshiba), it is assumed that users were relatively satisfied with both computers.

6.4 - Design Guidelines

The results of this study were used to produce product-specific guidelines. The rationale behind the guidelines will be provided, as well as implementation examples when possible. Guidelines will be listed in general categories. Although these guidelines are product-specific, and based on the results of this study, designers should not view these as a comprehensive list or final solution to prevent all usability issues. These guidelines should be used in conjunction with other information not available in this study, and reviewed judiciously along with other considerations to determine when the application of these guidelines is appropriate. Since the usability process is iterative, any design changes that result from the use of these guidelines should be subject to further usability testing. Future testing can help to confirm whether the inclusion of these product-specific design guidelines improve usability or not. The appropriate use of these guidelines should improve the usability of notebook computers among users with visual and upper extremity physical disabilities, as well as those without disabilities.

Keyboard Design

- Notebook computer keyboard design should be consistent with that of a desktop computer. Users with previous experience, limited vision or motor abilities may rely on highly learned motor skills to complete tasks, and learning the layout of a new keyboard may be difficult.
- Avoid placing keys such as the <Ins> key in locations traditionally reserved for the <Alt> key. The replacement of the <Ins> key where the <Alt> key is traditionally located may lead to catastrophic errors for users with total blindness when used in conjunction with the <F4> key (see Figure 6.1.4). Keys such as

the <Ins> key should retain a consistent configuration to that of a desktop computer keyboard (see Figure 6.1.2).

- Provide tactile identification such as Braille dots or raised bars on landmark keys, such as the <F>, <J>, <F4>, and <F8> keys. For the function keys, a separation between groups of keys may be preferred to the raised markings (Figure 6.4.1).

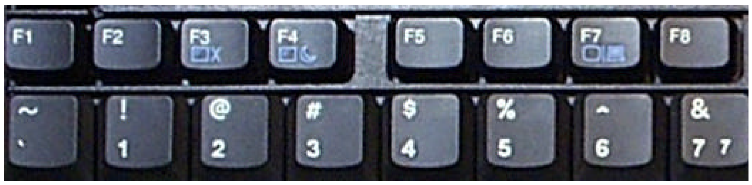


Figure 6.4.1 – Separation Between Function Key Groups on IBM Keyboard
(Note separation between <F4> and <F5>, as well as the separation between function keys and number keys.

- Keys that are larger in size than the character keys on a desktop keyboard, such as the <Caps Lock>, <Tab>, <Backspace>, etc. should be relatively larger in size than the character keys on the notebook computer keyboard. This difference in size will help touch typists and users with total blindness to differentiate them from character keys.
- Avoid placing characters that can be easily inserted into text in locations commonly used for the <Ctrl> and <Alt> keys (see Figure 6.1.1, and 6.1.3). If the position of the <Ctrl> and <Alt> keys is modified, users may need to correct errors resulting from failed shortcut attempts.
- Keys that are common on a desktop computer keyboard should not be omitted from the notebook computer keyboard. Please refer to Figures 6.2.1 and 6.2.2 for examples. Commonly used keys that are omitted may create difficulties, especially for experienced touch typists, users with visual disabilities, and those with limited use of one or both hands.

Pointing Device Design

- Provide an icon that leads to clear, detailed instructions on how to use the isometric pointing device on the Windows desktop. In some instances it may be appropriate to provide a multimedia tutorial to explain to users the different features of the device, how to program the device buttons, and possible interaction strategies that may increase user effectiveness. The desktop icon should be labeled with clear and precise wording, such as “Pointing Device Setup and Tutorial”. This should help users determine how the device can be adjusted to suit their preferences, which can help them become proficient with it more quickly.

- Pointing device buttons should be arranged in a layout that is familiar to users. Buttons that serve “left-click” and “right-click” functions should be arranged in a horizontal layout. The use of a horizontal layout is consistent with the layout on desktop computer mice, as well as the “left-click” and “right-click” metaphors used in many Windows-based applications. The buttons for the IBM pointing device provide an example of this horizontal layout (Figure 6.1.6).

Case / Display Design

- Case latches should be designed to allow for easy, one-handed operation, and not require fine motor skills to operate. The users in this study preferred one-handed operation. Also, users that may have limited use of either hand, or poor motor skills, may have difficulty operating latches that are not designed for one-handed operation. The latch for the Toshiba computer illustrated in Figure 6.2.3 provides an example.
- Buttons to open accessory devices, such as the CD and DVD drives, should provide enough protrusion for easy tactile identification and operation. Special care should be taken to ensure that the protrusion of these buttons is not such that they can be inadvertently activated. Inadvertent activation may interrupt the system and user, leading to errors and a loss of efficiency.
- PC card slots should be designed to assist a user with aligning the card in a single slot, if possible. A chamfered opening that assists the users with aligning the card may be an effective way of assisting users.
- PC card ejection mechanisms should allow for easy card ejection, and should eject the card at least an inch, if possible, to allow users to easily grasp and remove it. PC cards that are not ejected far enough may be difficult for users to grasp and remove.

- Glare should be minimized wherever possible. Users may not be able to comfortably view a display that is adjusted where glare can be minimized. Display surfaces should be designed to reduce glare as much as possible, regardless of the angle of the display and orientation to the light source. Excessive glare interferes with visibility and may lead to eyestrain (Sanders and McCormick, 1993).

User Interface Design Issues

- Icons used to denote system status should be easy to recognize, and easy to interpret. The user should not be required to remember complex metaphors and mentally transform an icon to determine its meaning. Figure 6.2.4 provides an example of an icon that is easier to recognize and interpret.
- Rollover text messages used with icons should be precise, specific, and provide a direct reference to the icon function. See discussion under "Perceptible Information" in section 6.2 for an example.
- User interfaces should provide specific, detailed, user-centered error messages, which help a user identify the specific problem, and how to solve it. Messages such as "Program X has caused a general protection fault and will be closed. If this problem persists, please restart the application or call technical support" are not written in user-centered terms, and do not guide the user to any specific resolution of the problem. An example of more specific, user-centered wording that provides a solution may be: "Program X could not open the remote document properly. Please save the document to a local drive by right clicking on the link, choosing 'Save target as' and naming the file. After the document has been saved locally, please try reopening it. "
- User interfaces should provide users feedback regarding status, and should allow users to proceed with, or cancel application updates if they are not desired. For example, when participants connected to the Internet using America Online, the application would try to obtain the latest update, without user approval. The application allocated a majority of the system resources to updating the application, possibly causing this and other applications to fail to work as normal.

Allow users to choose whether they want to update the application using dialog boxes containing messages such as “A new update is available for America Online. Would you like to update America Online to the latest version? <Update Now> <Cancel Update>”.

- User interfaces should provide a progress indicator for a software update, the estimated time remaining to update, and an option to allow the user to cancel the update at any time.

Table 6.4.1 - Universal Design and Usability Guidelines to Improve Notebook Computer Usability and Accessibility

Critical Incident:	Universal Design Principle:	Usability Principle:	IBM:	Toshiba:	Design Guidelines To Improve Usability and Accessibility
Keyboard not what users expect	Equitable use Flexibility in use Simple and intuitive use Tolerance for error	Learnability Memorability Errors Efficiency Satisfaction	Non-standard placement of <Fn> key.	Non-standard placement of several keys. Users unable to accomplish keyboard shortcuts with both hands. <Alt> key placement not consistent. Placement of <Ins> and <~> keys contribute to errors. Size of <Tab> and <Caps Lock> keys similar to other keys.	Place keys in locations and arrangements consistent with desktop keyboard. Do not omit keys that are common on desktop computers. Ensure special keys are the appropriate size.
Lack of feedback to user	Simple and intuitive use Perceptible information Tolerance for error	Efficiency Errors Satisfaction	Hardware, operating system, and application failures.	Hardware, operating system, and application failures. Battery level icon.	Provide specific, detailed feedback to correct the problem. Provide intuitive, descriptive icons with the proper use metaphor.
Pointing device difficulties	Equitable use Flexibility in use Tolerance for error Low physical effort	Learnability Efficiency Errors Satisfaction	Difficult to see Pointer difficult to use	Pointer difficult to use Buttons difficult to use	Improve contrast between pointing device and surroundings. Provide interactive software tutorial showing user pointing device features, button usage, and settings. Provide for easy tracking speed and sensitivity adjustment by placing “Pointing Device Properties” icon on desktop for users unfamiliar with Windows Control Panel settings.

Inadvertent key activation due to design	Equitable use Flexibility in use Simple and intuitive use Tolerance for error	Learnability Memorability Efficiency Errors Satisfaction	Difficulty completing <Ctrl> key shortcuts due to placement of <Fn> key.	Difficulty completing keyboard shortcuts due to non-standard placement of several keys. Users unable to complete <Ctrl-key> shortcuts using both hands. <Ins> and <~> keys located where <Alt> key expected. Non-standard key placement contributes to errors.	Place keys in locations and arrangements consistent with desktop keyboard. Do not omit keys that are common on desktop computers. Ensure special keys are the appropriate size
System state not apparent to user	Equitable use Simple and intuitive use Perceptible information	Efficiency Errors Satisfaction	Hardware, operating system, and application failures.	Hardware, operating system, and application failures.	Provide detailed information regarding system state. Do not change system state without warning user.
Does not guide user to a solution	Equitable use Simple and intuitive use Perceptible information	Efficiency Errors Satisfaction	Hardware, operating system, and application failures.	Hardware, operating system, and application failures.	Provide specific, descriptive feedback to guide the user to a solution.
User cannot accomplish tasks using chosen method	Equitable use Flexibility in use	Efficiency Errors Satisfaction		Omission of right <Ctrl> key.	Use the common QWERTY keyboard layout.
Case design	Equitable use Flexibility in use Tolerance for error	Efficiency Errors Satisfaction	Difficult to open case with one hand. CD drive easy to open inadvertently.		Provide a design that allows for one-handed operation. Design button on CD drive with enough protrusion to easily feel, but not so much that it is frequently pressed inadvertently.

6.5 - Lessons Learned

Limitations to data collection, experimental methods, computer reliability, and participant motivation may reduce the ability to generalize these results. Although these issues cannot always be eliminated, researchers should be aware of these issues and work to minimize them in the future. The intent of this study was to obtain a rough characterization of usability and accessibility issues over a short period of time, and was not intended to replace other usability evaluation methods.

This study was intended to collect usability data by employing a remote ethnographic method and critical incident recording technique, with supplemental data collected using a questionnaire. Although participants recorded several critical incidents, the number recorded was much less than expected. Critical incident data recording may have been greater had participants had more formal instruction on critical incidents, and a different definition of a critical incident. Due to the exploratory nature of this study, an open-ended definition of critical incidents was used ("any interaction with the hardware or user interface that is surprisingly positive or surprisingly negative") to try to elicit as much information as possible, without leading participants. Participants may have been unsure as to whether an incident was severe enough to report, thus did not report many critical incidents. When asked about the low critical incident reporting rate, one participant responded "I guess there were many instances where a critical incident occurred, however, I am used to those problems, so I didn't report them." The definition of critical incidents should have included "any interaction with the interface that is bothersome, irritating, difficult, as well as any interactions that were helpful or pleasing." A definition such as this may have increased the reporting rate.

Formal instruction on critical incident reporting was contemplated, however it was decided against due to logistical reasons. Since most participants in the notebook computer study also participated in the cellular telephone study, it was believed that the time to complete the cellular telephone study, and formal instruction on critical incidents would be more than participants would allow. Due to time, scheduling difficulties, and transportation logistics, it was believed that participants would not complete the study if multiple study sessions were required. Formal instruction on critical incidents may have provided participants examples of what could be considered a critical incident, and increased the reporting rate.

The remote ethnographic method used in this study was intended to capture context-of-use data through critical incident reporting technique. Although software such as Camtasia can record critical incidents and capture issues internally, the remote ethnographic method does not allow a researcher to capture important external context-of-use information. External context of use information can include information on the locations where the computer is used, method of use, peripherals connected, and daily routines (O'Brien, Rodden, Rouncefield and Hughes, 1999). Use locations for a portable device such as this can include the couch, kitchen table, desk, etc. Method of use information can include whether the device was used on a supported surface (table or desk) on an unsupported surface (using on his or her lap), the different postures observed during use (use while sitting up, laying down), etc. Information on peripheral connectivity can include connections to printers, accessibility devices, or even the ability to conveniently access a telephone jack to connect to the Internet. Daily routines can include information on when the computer is used, what is it used for, is it used before work, after work, before bedtime, and how the computer may be interwoven into these routines. The remote ethnographic method unfortunately does not allow this external context-of-use information to be captured. According to O'Brien et al. (1999) this information is often important for designers to fully understand how a product is used by different types of users in different environments, and may generate new ideas to solve issues that were observed (i.e. a wireless adaptor to allow use of the Internet without being directly connected to a telephone jack). For those with disabilities, this information may provide ideas on universal design features that can improve accessibility, or information on further considerations for the design of notebook computers.

Low critical incident reporting and QUIS results may have been affected by previous computer experience, and the availability of a desktop computer. Participants in this study were required to have experience with personal computers, which was intended to provide participants with a base knowledge of computer use to compare the use of the notebook computer to. As participants started to experience difficulty with adjusting to two computers (desktop computer and notebook computer), many participants curtailed the use of the notebook computer, and returned to using their desktop computers since those machines were already configured for their needs. In many cases, difficulties resulting in low usage were not even related to the notebook computers. Participants noted trouble with installing software to transfer files from their desktop computer, lack of passwords required to transfer software, or unwillingness to spend time obtaining readily available, no-cost technical support to connect the computer to a network as reasons for limited use. Once participants lost motivation due to these or other issues, they did not use the notebook computer as frequently. If a participant lost motivation in the first loan period, they may not have regained it in the second loan period, which may have affected the critical incident reporting rates and the QUIS results.

A within-subjects study design may have been less effective for collecting usability data than a between-subjects study design. Since a within-subjects study design uses the same participants for each level of treatment, the amount of data collected may have been reduced due to learning effects, and lowered participant motivation. Once participants gained experience using one notebook computer and recorded critical incidents, a learning effect may have occurred and participants either determined the solution to the same critical incidents on the new computer, or may have expected the same issues to occur and did not report them. Additionally, once a participant lost motivation due to a technical issue, they did not regain motivation when presented with the second notebook computer, resulting in non-equivalent computer use, poor utilization of the computer, and poor utilization of study time. A between-subjects study design may have improved data collection, and the utilization of valuable resources. A between-subjects study design was contemplated, however due to practical constraints such as a limited pool of eligible participants and limited study resources, a within-subjects study design was used.

The length of the computer loan periods may have been too long to collect data efficiently. Participants commented that they used the computer frequently at first, but then the usage dropped off considerably shortly thereafter, and computers sometimes were unused for long periods of time. The reduction in usage may be attributed to technical problems discussed earlier, which may have resulted in a lack of participant motivation. The lack of usage results in very low equipment utilization, and due to the expense of the equipment used in this study, efficient utilization is especially important. In an applied setting where resources are limited and deadlines are critical, a shorter study period among more participants may allow for better utilization of equipment and better data collection within a shorter period of time. The use of participants from an academic setting created further complications, since it was very difficult to schedule a 60-day study period that did not overlap with a school holiday or critical exam periods. Participants in this type of environment are often not available at these times, thus reducing the ability to efficiently exchange computers, resolve technical issues, etc. Participants in this environment, while an important user demographic, may also lack the maturity to follow through with commitments to the experiment.

6.6 - Future Study Opportunities

The results of this studied provided specific information on the usability of notebook computers among users with visual and upper extremity physical disabilities. Although this information is helpful, it also revealed many other study opportunities that may be of future interest. These opportunities include repeating the study with a between-subjects design, using participants with no experience, using a shorter loan period, testing the usability of icons and icon metaphors, and repeating the experiment with users that are formally trained in the critical incident reporting technique.

Between-Subjects Study Design

Due to the limited population of people with visual and upper extremity physical disabilities in the local area, a within-subjects study design was required. The within-subjects study design may have been a factor in the low rate of critical incident reporting. Once participants experienced a critical incident and reported it while using the first model of computer, they may have experienced a similar critical incident and failed to report it when using the second computer. This potential gap in reporting across computers may have been due to a learning effect with the first computer, where the participant learned how to prevent a critical incident, or the participant may have expected the critical incident and since it was not new, neglected to report it. The use of a between-subjects design may alleviate these issues, and lead to improved data collection.

Participants with No Computer Experience

The use of participants with previous experience may have reduced the critical incident reporting rate. Participants in this study were required to have computer experience, however, many had limited notebook computer experience. Participants with previous experience may expect usability issues to occur at specific times, or may know ways to circumvent them, which may affect the reporting rate.

Since the participants had previous computer experience, novice computer users were not represented in this study. The use of novice users may have revealed important usability issues that may not occur with users with previous experience. Novice users may be difficult to locate, however, due to the pervasiveness of technology in society, and the increasing reliance of it for daily activities for those with disabilities. Regardless, the use of experienced and inexperienced users is important in any technology study, and novice computer users were unfortunately under represented in this study.

Notebook Computer as Primary Computing Platform

The participants in this study all owned, or had access to a desktop computer that was configured for their needs. Since participants had an already configured computer readily available, the notebook computer may have been used in a recreational, or exploratory manner, and not to accomplish important work.

The use of notebook computers as a primary computing platform may be increasing due to the portability of these devices, decreasing cost, and business or academic requirements. Due to this, it is especially important that the usability of these devices as a primary computing platform be studied, since users may not have a desktop computer available, and the postures required to use notebook computers may not be comfortable, or even possible for those with disabilities (Sommerich, 2002). The study of a notebook computer as a primary computing platform may also alleviate low usage issues where participants could use another computer to accomplish work when a problem occurred, rather than resolving the problem and continuing work with the notebook computer.

Shorter Loan Period

The loan periods in this study were each 30 days in duration. While the use of 30-day loan periods allowed for comprehensive use of the computer by some, and intermittent use by others with busy schedules, it may have been less effective than using a shorter loan period.

The use of a 15-day loan period may actually allow for increased data collection than the use of a 30-day loan period. Users in this study were generally very motivated to participate in the first several days of the experiment. As the experiment progressed, participant motivation waned, and the computers went unused. Since the computers went unused, no data was being collected. Rather than allow this expensive equipment to go unused by continuing the loan periods, the computers could have been reallocated to new participants, who could have provided additional data.

Although a 15-day loan period is suggested, it is still important to obtain data on long-term usage. Participants that use the notebook computer as their primary computing platform should continue to use the 30-day loan periods, while those that reserve the notebook computer for secondary use should be limited to 15-days.

Usability of Icons and Metaphors

The battery level icon on the Toshiba computer seemed problematic for participants that did not use the AC adaptor all the time. Since the battery level icon was not apparent, the computer may have shut down on participants inadvertently.

A suggested design for a new battery level icon was provided earlier in the text. This design, as well as other icon and metaphor designs for battery level indicators should be tested. It would be helpful to determine the extent of customer support costs that may have been incurred due to problems associated with the original icon design (the light bulb), and how much customer support costs would be reduced due to a more usable redesign.

Repeat Experiment with Formal Critical Incident Training

The definition of a "critical incident" used in this study was "any interaction with the hardware or user interface that was surprisingly positive, or surprisingly negative." Since this study was exploratory in nature, participants were provided with a very broad definition with the expectation that several critical incidents would be reported. While some participants did report critical incidents, the reporting rate was much less than what was anticipated. Since participants were required to have previous computer experience, the low reporting rate may have been due to participant experience, or the lack of examples of what could be considered a critical incident. This experiment should be repeated with new participants that have received formal critical incident training to determine if additional usability issues exist that were not reported.

Use of a Semi-Remote Ethnographic Method

The use of a remote ethnographic method in this study did not capture important external context-of-use information that may be vital in determining usability issues, since a researcher was not available to spend time in each participant's home. Future studies can try a "semi-remote" ethnographic method, where internal context-of-use data is collected through the use of Camtasia, and external context-of-use data is collected via periodic evening visits to participant homes. When visiting participant's homes researchers can make note of computer placement, method of use, peripheral connectivity, and possibly even daily routines. Obrien et al. (1999) used a similar method (evening visits) in a study of a system to deliver broadband communication services into the home.

Use of Desktop Video Recording Software to Record Critical Incidents

Camtasia is a desktop video recording application that allows users to record screen activity while providing a running narrative. One of the benefits of using applications such as this is that user interface researchers can see exactly what users are trying to accomplish and what actions occurred, while listening to the user's narrative. Although the software allows for robust data collection if this method is used, users in this study did not record critical incidents in this manner.

At the beginning of the study, participants were asked to record critical incidents by activating Camtasia, reenacting the critical incident, providing a narrative, and then saving the recording. The experimenter demonstrated this technique to each participant prior to the first computer loan period. Only one of the eight participants, however, followed these instructions and reported critical incidents using this method. Instead, participants used a retrospective think-aloud technique and only recorded the narrative associated with the critical incidents. Future study opportunities can investigate whether methods to use this tool for collecting usability data (such as embedding it in an application) can increase effectiveness, as well as why some participants elect not to reenact critical incidents and only provide the narrative.

7.0 – Notebook Computer Study Conclusions

The usability of notebook computers was studied among users with visual and upper extremity physical disabilities using a remote ethnographic method with a critical incident reporting technique. Participants were asked what their computer preference was before the study, and after the study was complete. Supplemental data was collected using a modified version of the Questionnaire for User Interface Satisfaction (QUIS 7.0). Participants included those with visual and upper extremity physical disabilities.

Critical incidents were reported for both computers. Keyboard layout, pointing device use, feedback from the user interface, case design, and display glare were among issues reported. The Toshiba computer received the most critical incident reports due to keyboard layout, while the IBM received the most reports due to case design.

Supplemental data revealed very little difference in preference between the computers. Overall results for the QUIS survey were more favorable for the IBM, however differences in QUIS results were between the computers were not significant. Before and after preferences indicated that most participants that preferred one computer before the study began, preferred the same computer at the conclusion of the study.

The results of the critical incident reports were analyzed to determine where universal design principles and usability principles were either demonstrated or violated. Product-specific design guidelines were created to help designers adhere to these principles in future designs. The use of product-specific guidelines and the inclusion of universal design features can help to increase the usability and accessibility of notebook computers, regardless of disability. Manufacturers that incorporate these features into products can enjoy greater market share, while organizations that adopt products that include these features can comply with accessibility-related regulations, and provide a consistent set of computing tools to employees, regardless of disability.

The use of a remote ethnographic method with a critical incident reporting technique allows researchers to obtain data on relatively critical usability issues within a short period of time. Although this method may have revealed severe usability issues, it is not intended to replace other usability inspection methods, and should be used in conjunction with them to determine the extent of usability issues.

The QUIS allows researchers to collect user satisfaction data using a broad variety of categories. Although this tool may be very helpful for software-based user interface evaluations, it was troublesome to use for hardware-based evaluations. This tool may be helpful if extensively redesigned to include specific, non-biased hardware-related questions. Even with these questions, the QUIS should be used in conjunction with other usability methods to determine what usability issues exist.

8.0 - Experimental Methods - Cellular Telephone Study

8.1 - Overview

The previous study by Smith-Jackson et al. (2001), revealed that cellular telephone users with visual and upper extremity disabilities have concerns about the design of the keypad, the display, and the auditory feedback. To evaluate the two former concerns, several cellular telephone prototypes were used, each with slight modifications to the display or the keypad. This study was designed to evaluate which display and keypad attributes provide the best performance and highest user satisfaction among those with visual and upper extremity physical disabilities. Auditory feedback was not evaluated, since prototypes with auditory displays could not be provided.

8.2 - Design

A mixed factors study design was used, with lateral pitch (Figure 8.2.1), key height (Figure 8.2.2), keystroke (Figure 8.2.2), location of the <Power>, <Send>, and <End> keys, and numeric and alphabetic font size (Figure 8.2.3), assigned as within subject independent variables (Table 8.2.1). The between subjects independent variable was disability type (Table 8.2.1).

Table 8.2.1 - Independent Variables

Independent Variable:	**Factor Type:**	**Number of Levels:**	**Levels:**
Lateral Pitch	Within	4	10 mm, 11 mm, 12 mm, 13 mm
Key Height	Within	3	0.3 mm, 0.5 mm, 0.7 mm
Key Stroke	Within	3	0.3 mm, 0.5 mm, 0.7 mm
Numeric Font Size	Within	6	16, 22, 36 point font, each in normal and bold weights
Alphabetic Font Size	Within	3	16, 22, and 36 point font
Power key locations	Within	8	Refer to Table 2 for a description of each model
Send / End key locations	Within	8	Refer to Table 2 for a description of each model
Disability Type	Between	5	Total blindness, legal blindness, minor and severe upper extremity disabilities, no disability

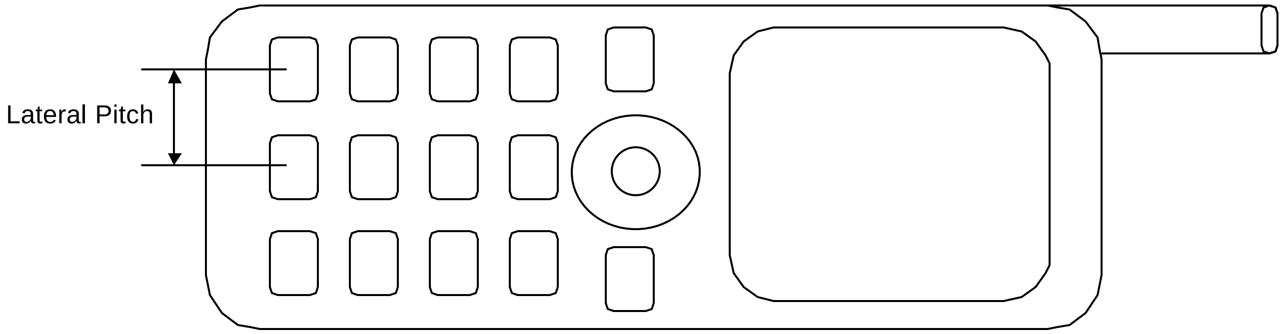


Figure 8.2.1 - Illustration of Lateral Pitch

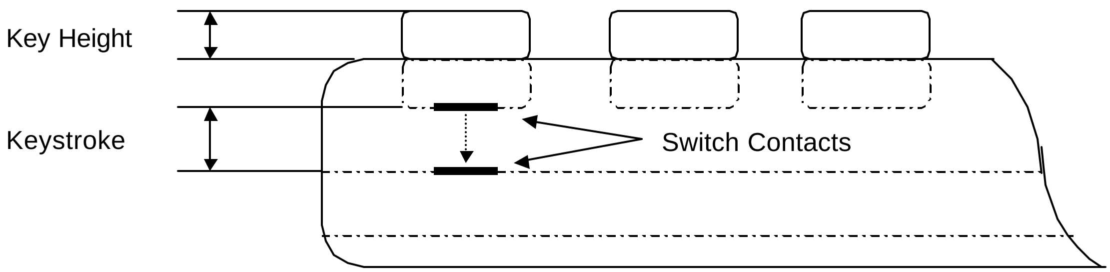


Figure 8.2.2 - Illustration of Keystroke and Key Height (Exploded Side View)

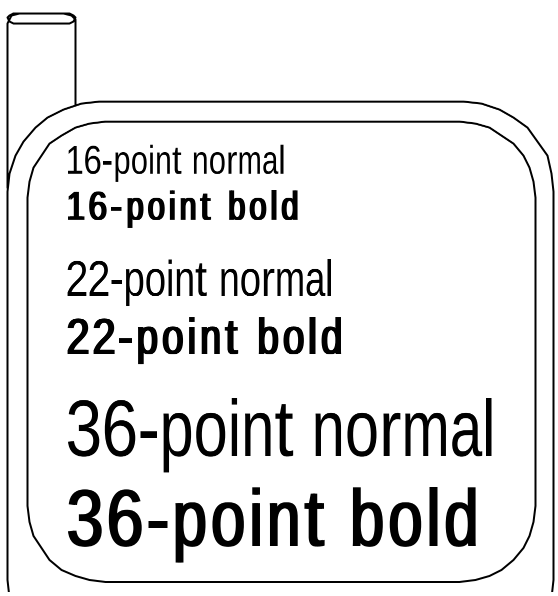


Figure 8.2.3 - Illustration of Display Font Sizes
(Example fonts are not actual size)

Participants completed several common cellular telephone dialing tasks during the study. Dialing tasks were chosen since participants in the Smith-Jackson et al. (2001) study indicated that they generally make outgoing calls, and did not receive them. Additionally, since the cellular service was not activated, user performance while receiving calls could not be measured.

Task completion times, the number of failures for each task, and subjective ratings of the attributes served as the dependent variables. The time to complete the task was measured to the nearest tenth of second from when the participant pressed the first key, until the participant pressed the <Send> key. Dialing tasks were treated as Boolean operations with only 2 outcomes – success or failure. Success occurred when all the required numbers were dialed correctly, and in the proper sequence. Failure occurred when the required numbers were dialed incorrectly or out of sequence. A dialing task may include several distinct dialing errors, which overall will result in a task failure. Task failures were quantified instead of dialing errors for simplicity of analysis, and due to the unavailability of electronic instrumentation that could detect minor errors. When a task failure occurred, the task completion time was reset, and the trial repeated. Only times for successful task completion were recorded and used in the analyses.

After completing the dialing tasks, participants rank ordered their preferences for each level of independent variable, based solely on the independent variable of interest (i.e. lateral pitch, keystroke, etc.). Due to the limited availability of cellular telephone models, the levels of lateral pitch were presented using four different models, rather than four different lateral pitches in the same model. To reduce the possibility of selection bias that could result from participants selecting a model based on other variables, every effort was made to isolate each independent variable as best as possible. Although the potential biases due to differences in models can limit the ability to generalize the results, these biases cannot always be avoided.

After participants ranked their preferences for each independent variable, they were asked why they chose each particular model. If a participant chose a model based on a variable other than the independent variable of interest, that portion of the experiment was repeated, again emphasizing the importance of ranking each model based solely on the independent variable of interest.

Participants were then presented with each level of independent variable, and asked to rate it on “Ease of Use” and “Accessibility”. “Ease of Use” was defined as “how easy this specific feature (independent variable) is to use”. “Accessibility” was defined as “the operational suitability of a specific feature (independent variable) to accomplish the task” as well as “the ability to get to and use that feature”. Ratings were recorded on a 7-point Likert-type scale.

8.3 - Participants

Fifteen people were recruited from the local area to participate (Table 8.3.1). Participants included those with legal blindness (LB) and total blindness (TB), minor and severe upper extremity disabilities (MU and SU, respectively), and those with no apparent disabilities (ND). All participants were required to have at least minimal use of both hands. It was not feasible to include those without use of their hands or arms, since operation of the telephone required that the device be held and manipulated with both hands. Participants were categorized into disability groups based on self-reports, with three participants in each group.

Previous cellular telephone experience was not required for participation. Due to the nature of the tasks (dialing a telephone then providing subjective measures of specific features), previous experience may have had little benefit, and potentially biased participants toward models similar to their own. Also, due to the limited study population in the area, finding participants with previous cellular experience may have further complicated recruiting efforts.

Table 8.3.1 - Cellular Telephone Study Participants

#:	Disability Group:	Age:	Gender:	Disability Description:
1	TB	66	M	Total blindness
2	ND	23	F	No disability
3	MU	35	M	Chronic numbness in right hand and arm (minor upper extremity disability)
4	TB	41	F	Total blindness
5	LB	19	M	Legal blindness
6	LB	25	F	Legal blindness
7	ND	30	F	No disability
8	LB	22	F	Legal blindness
9	TB	47	F	Total blindness
10	MU	19	M	Reduced fine motor control in hands (minor upper extremity disability)
11	SU	20	M	Cerebral palsy, very limited use of left hand (severe upper extremity disability)
12	ND	21	F	No disability
13	SU	58	M	Parkinson's disease - severe hand and arm tremors (severe upper extremity disability)
14	SU	52	F	Cerebral palsy – arthritis in both hands / wrists, limited use of both hands and upper body (severe upper extremity disability)
15	MU	26	M	Minor hand tremors when performing motor tasks (minor upper extremity disability)

8.4 - Apparatus

Dialing tasks were monitored throughout the experiment using a video camera, with the output directed to a VCR. The camera was mounted overhead in a vertical position to record hand movements and key presses.

Eight cellular telephone models were used in the dialing task portion of this experiment. Audiovox models CDM-4000, CDM-9000, and CDM-9100 were tested, as well as Toshiba model CDM-310T (Table 8.6.1, models 6, 8, 1, and 9 respectively). Each model was either a current production version, or had minor modifications. Each CDM-9000 differed in keystroke, each CDM-9100 differed in the amount of key protrusion, and models CDM-310T, CDM-4000, CDM-9000 and CDM-9100 each differed in lateral pitch. Several additional models of CDM-9100 were used to determine participant font preferences, while models 1-8 (Table 8.6.1) were used to determine preferences for <Power>, <Send>, and <End> key locations. These models were chosen because they are representative of contemporary cellular telephone models, and were easily produced with different levels of the independent variables.

Each model was labeled with an alphanumeric code to ensure that the proper model was presented to the participant in the appropriate order. An example is the label “B-1” for the model CDM-9000 with a keystroke of 0.3 mm. The label allowed for easy identification of models with specific features, and was not shown to participants.

8.5 - Experimental Hypotheses

The experimental hypotheses for this study are listed below:

Hypothesis 1: There will be main effects and interaction effects of key attributes and disability on task performance and subjective evaluations.

Hypothesis 2: There will be main effects and interaction effects of disability and display font size on subjective evaluations.

Hypothesis 3: There will be main effects and interaction effects of disability and “Power”, “Send”, and “End” key locations on subjective evaluations.

8.6 - Procedure

Each participant was asked to read and sign an Informed Consent form approved by the Virginia Tech Institutional Review Board prior to the start of the study (Appendix C). Participants with visual impairments participants had the form read to them.

Each participant was given experimental instructions via audio recording to ensure that participants received a consistent set of instructions (Martin, 2000). Participants were then asked to complete several dialing tasks using the models 1, 6, 8, and 9 (Table 8.6.1). These dialing tasks allowed each participant to become familiar with the types of experimental tasks, and the use of each telephone model.

To reduce the confounding effects of learning and practice, a partially counter-balanced Latin square was used to determine the presentation order of each model. Since there were 15 participants, and 12 model presentations, a fully counter-balanced study was not possible

Each experiment involved performing tasks common to cellular telephone use, including:

1. Dialing a 7-digit number chosen at random.
2. Dialing an 11-digit number chosen at random.
3. Dialing a number from the recall memory chosen at random.

4. Dialing “911” (Used to call for help in an emergency in the United States).

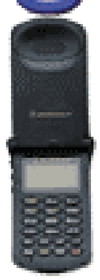
To reduce the effects of inter-trial variability for each type of task, participants completed the 7 and 11-digit dialing tasks five times, and the recall and emergency call tasks three times. The completion times and number of failures for each set of tasks were used in the data analysis.

Participants dialed the telephones by holding the telephone with one hand, and dialing with the other. Although some participants may not have preferred this dialing method, it helped to provide a consistent dialing method between disability groups. Participants with visual and physical disabilities may not be able to dial the telephone with one hand using the thumb, as those with no disabilities often do. Additionally, due to the large number of trials in the experiment (256), it was believed that participants would become fatigued and unable to complete the dialing tasks in a consistent manner if the use of a one-handed dialing method was allowed. Once participants began the dialing portion of experiment, they were not allowed to switch hands or use an alternate dialing method until the dialing portion of the experiment was complete. In contrast, participants were allowed to dial the telephone however they wished during the subjective evaluation portion of the experiment.

After the dialing tasks were completed, models were grouped based on the independent variable of interest (i.e. keystroke, lateral pitch, etc.). Participants were provided with a subjective evaluation form (see Appendix D) in which they were asked to rank order each model in order of preference, based solely on the independent variable of interest (Table 8.2.1). Lateral pitch preferences were determined using one of each of the four basic models. Keystroke preferences were determined using the CDM-9000 models. Key height preferences were determined using the CDM-9100 models. Font size preferences were determined using a separate set of CDM-9100 models programmed specifically for this portion of the experiment. <Power>, <Send>, and <End> key preferences were determined using several different telephone models (Table 8.6.1).

Models 1-8 (Table 8.6.1) were used in the <Power>, <Send>, and <End> key evaluations, however, model 9 was not. This model was excluded from these evaluations since the <Power>, <Send>, and <End> keys were similar to other models used.

Table 8.6.1 –Cellular Telephone Study Models

#:	Phone:	Model:	Location of <Power> Key	Location of <Send> / <End> Keys
1		Audiovox CDM-9100 11 mm lateral pitch model used in dialing tasks. Key height levels of 0.3 mm, 0.5 mm, and 0.7 mm used in dialing tasks.	Next to circular function key, above <Clr> key, and below <II> key.	Adjacent to circular function key, above <V+> and <Clr> keys and below <I> and <II> keys.
2		Nokia 8260 Not used in dialing tasks.	Very small key on top right edge of telephone.	Silver horseshoe shaped key above numeric keypad, below display.
3		Ericsson T28 World Not used in dialing tasks.	Scalloped shaped key above numeric keypad. Top edge of key is molded into the display face. Labeled “No”.	Scalloped shaped keys above keypad, along right and left edges of bottom of display. Top edge of key is molded into the display face.
4		Motorola V2260 Not used in dialing tasks.	Round key located on lower left hand corner of telephone.	Round keys located above other keys. <Send> and <End> key order is reverse of other models.
5		Motorola Star-Tac Not used in dialing tasks.	Key located on lower left corner of keypad.	Keys located on lower right hand side of keypad, one above the other.

6		Audiovox CDM-4000 13 mm lateral pitch model used in dialing tasks.	Highest key on right side of keypad. Above circular multifunction key, below display.	Highest keys on keypad. Above circular multifunction key, below display.
7		Sanyo SCP-5000 Not used in dialing tasks.	Key above the <3> key, on lower right hand side of multifunction key.	Keys above <1> and <3> keys, to the lower left and lower right of multifunction key.
8		Audiovox CDM-9000 12 mm lateral pitch model used in dialing tasks Keystroke levels of 0.3 mm, 0.5 mm, and 0.7 mm used in dialing tasks.	Adjacent to circular function key, above <3> key, and below <Clr> key.	Adjacent to circular function key, above <1> and <3> keys and below "<Sto> and <Clr> keys.
9		Toshiba CDM-310T 10 mm lateral pitch model used in dialing tasks. Not used in <Power>, <Send>, and <End> key evaluations.	Located above <3> key, and below <Clr> key. Same size, shape, and separation as numeric keys.	Located above the <1> and <3> keys. Same size, shape, and separation as the numeric keys.

After participants ranked each model based on the independent variable of interest, they rated each level of independent variable based on ease of use, and accessibility. Ratings were conducted using a 7-point, Likert-type scale (see Appendix D). Unlike the ranking portion of the experiment, each level of independent variable was rated on ease of use and accessibility independent of all other levels.

<u>8.7 - Analysis of Cellular Telephone Study Data</u>

Results for the dialing tasks, rankings, and ratings were analyzed using a series of two-way analysis of variance (ANOVA) models, where the independent factors were disability, cellular telephone attribute (keypad or display variables), and the disability by attribute interaction (Table 8.2.1). Subsequent one-way ANOVAs were run for the aggregate of all disability categories (Overall), and for each specific disability category (LB, TB, MU, SU, ND) to determine specific differences in objective and subjective measures for each disability and level of independent variable. Tukey Honestly Significant Difference (Tukey-HSD) tests were used when an overall significant difference was present.

Analyses were conducted with a significance level of 0.10 ($\alpha = 0.10$) to minimize the Type I error, or likelihood of rejection of the null hypotheses when the null hypotheses are true. The null hypothesis was rejected if the probability of a chance outcome ("p" value) was less than or equal to the significance level (α). The significance level of 0.1, rather than 0.05, was chosen to increase power, since the sample size is relatively small, and to increase the likelihood of finding significant effects that might warrant further investigation.

Results are reported for the two-way ANOVA first, then by the one-way ANOVA. Overall significant differences revealed when using the one-way ANOVAs and Tukey-HSD tests are indicated on the graphs through the use of labeled groups (Group A, B, C, etc.). In several instances the results of two-way ANOVAs revealed significant main effects and interactions, but subsequent one-way ANOVAs did not. In cases where the independent variables were not significant, the level that produced the best outcome (e.g. performance time or number of task failures) is reported.

Results from each disability group (LB, TB, MU, SU, and ND) are collapsed into specific categories to simplify data reporting when possible. Results for the “Visual” category consist of the mean measures of participants with legal and total blindness. Results for the “Physical” category consist of the mean measures of participants with minor and severe upper extremity disabilities. Results for the “None” category consist of the mean measures of participants with no disabilities. Finally, results for the “Overall” category consist of the mean measures across all participants.

9.0 - Results - Cellular Telephone Study

Results of all analyses are reported below, organized by the specific attribute of interest. Reported figures are sorted in the order of best performance, using the Overall category as the sort criteria. Therefore, the sort order on the X-axis differs in several of the figures, which allows for easier comparison across disability categories.

When a significant difference is determined for the “Overall” measure, the levels of the independent variable that are not significantly different from one another are grouped and denoted by a line spanning the appropriate levels over the X-axis. Levels of the independent variable in separate groups (i.e. Group A, B, C) are significantly different.

9.1 - Effects of Lateral Pitch

7-Digit Dialing Tasks

A two-way analysis revealed significant effects on the completion times for disability ($F_{(4,29)} = 7.26$, $p < 0.005$), lateral pitch ($F_{(3,29)} = 24.88$, $p < 0.0001$), and the lateral pitch by disability interaction ($F_{(12,29)} = 1.99$, $p < 0.02$). A subsequent one-way analysis revealed a significant difference in task completion times between the 12 and 13 mm lateral pitch groups, when compared to the 10 mm and 11 mm groups overall ($p < 0.001$). Participants in the LB, MU, and ND participant groups, performed the fastest using the 12 mm lateral pitch. Participants in the TB and SU groups performed the fastest using the 13 mm lateral pitch. One-way analyses for each group did not reveal a significant difference in task completion times between the 12 mm and 13 mm levels of lateral pitch (Figure 9.1.1).

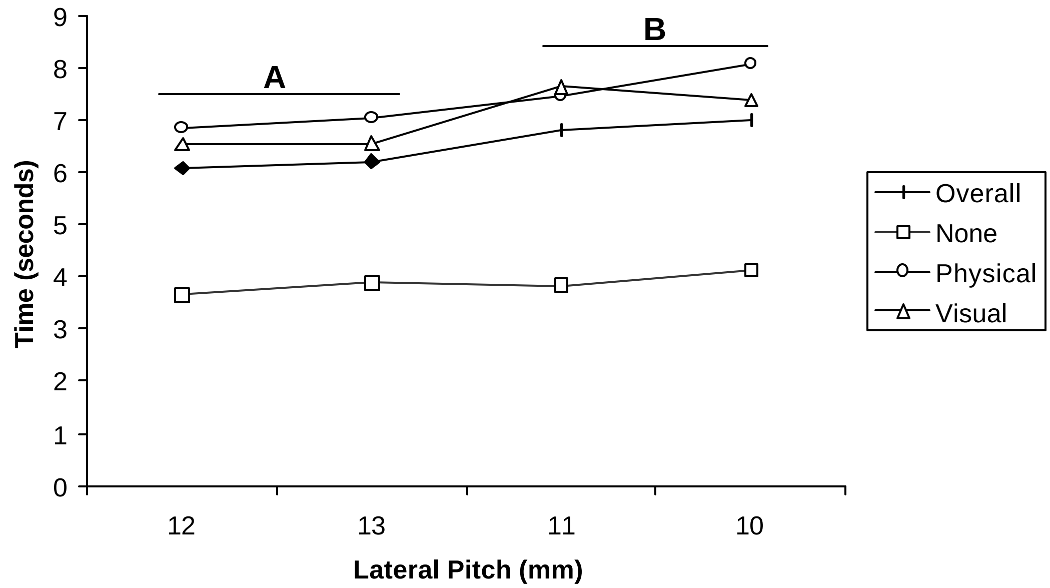


Figure 9.1.1 - Mean 7-Digit Dialing Time by Lateral Pitch and Disability
(Groups A and B are significantly different overall)

A two-way analysis revealed significant effects on the number of task failures for disability ($F_{(4,29)} = 8.47$, $p < 0.003$), lateral pitch ($F_{(3,29)} = 3.46$, $p < 0.016$), and the lateral pitch by disability interaction ($F_{(12,29)} = 5.96$, $p < 0.0001$). A one-way analysis revealed a significant difference in the number of task failures between the 10 mm and 12 mm levels of lateral pitch, with the lowest number of task failures occurring with the 12 mm lateral pitch ($p < 0.05$). All participant groups, except the MU group, exhibited a significant difference in task failures due to lateral pitch, with each group committing the fewest task failures using different levels of lateral pitch. Participants in the MU group did not commit any task failures during the study (Figure 9.1.2).

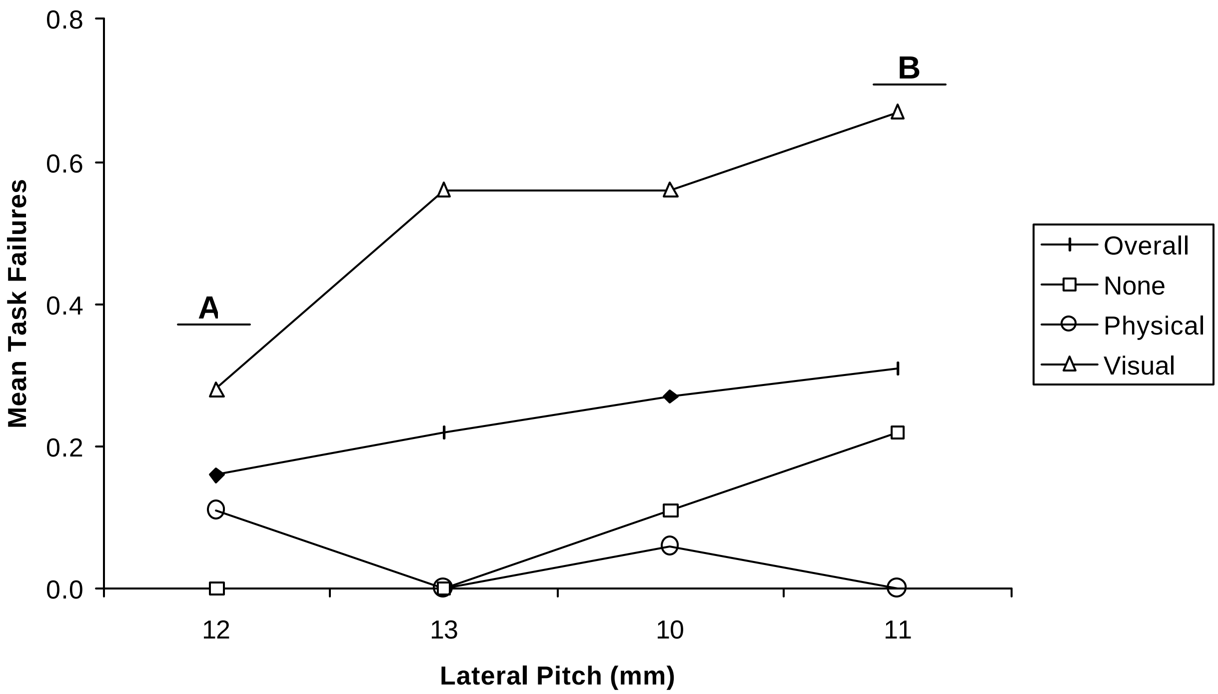


Figure 9.1.2 - Mean Task Failures for 7-Digit Dialing Tasks by Lateral Pitch and Disability
(Groups A and B are significantly different overall.)

11-Digit Dialing Tasks

A two-way analysis revealed significant effects on task completion times for disability ($F_{(4,29)} = 7.71$, $p < 0.004$), lateral pitch ($F_{(3,29)} = 21.50$, $p < 0.0001$), and the lateral pitch by disability interaction ($F_{(12,29)} = 2.22$, $p < 0.009$). One-way analysis, overall, revealed a significant difference in task completion times between the 10 mm and 12 mm levels of lateral pitch ($p < 0.03$), with the fastest task completion times occurring with the 12 mm lateral pitch. The TB, MU, and LB participant groups performed the fastest using the 12 mm lateral pitch. The ND and SU groups, however, performed the fastest when using the 13 mm lateral pitch (Figure 9.1.3).

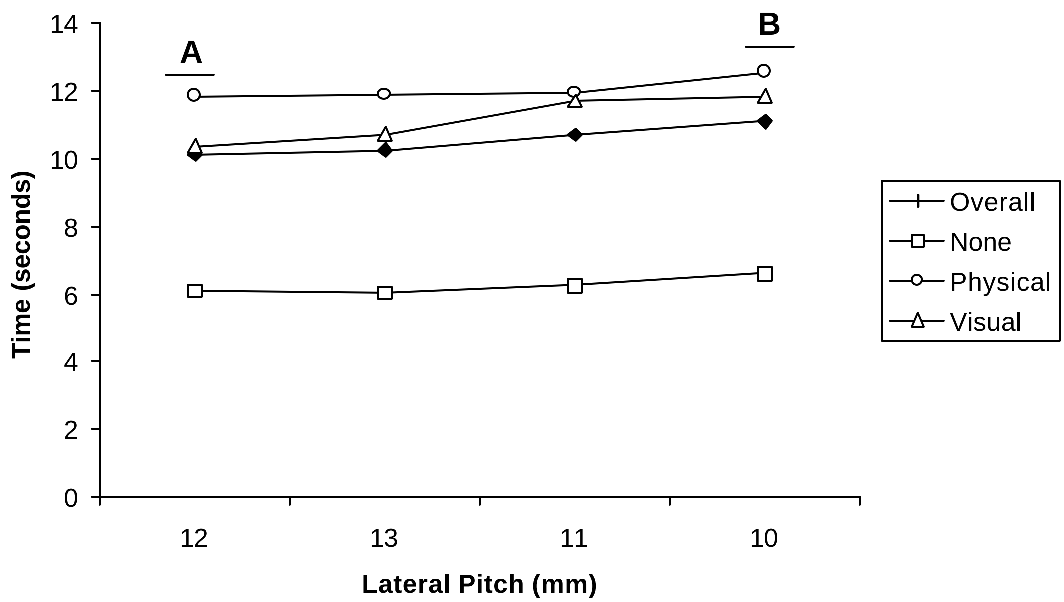


Figure 9.1.3 - Mean 11-Digit Dialing Time by Lateral Pitch and Disability
(Groups A and B are significantly different overall.)

A two-way analysis revealed significant effects on the number of task failures for disability ($F_{(4,29)} = 27.09$, $p < 0.0001$), lateral pitch ($F_{(3,29)} = 19.56$, $p < 0.0001$), and the lateral pitch by disability interaction ($F_{(12,29)} = 13.07$, $p < 0.0001$). One-way analyses, overall, revealed a significant difference in the number of task failures between the 12 mm and 13 mm levels of lateral pitch, when compared to the 10 mm and 11 mm levels ($p < 0.0001$) (Figure 9.1.4). The Visual and None categories exhibited the fewest task failures when using the 12 mm lateral pitch. Participants in the Physical category had the lowest number of task failures when using the 11 mm lateral pitch (Figure 9.1.4).

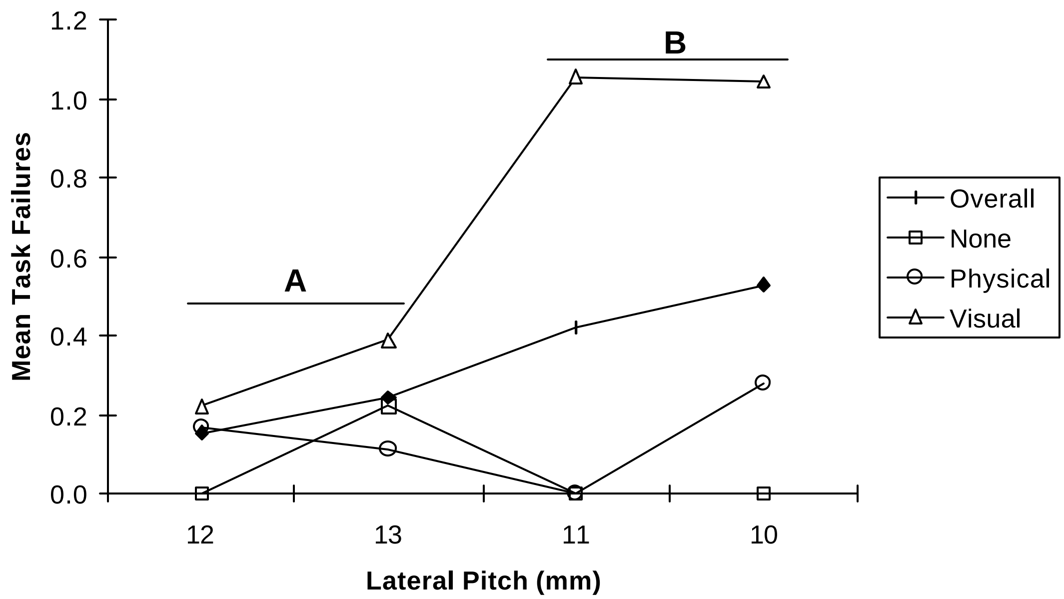


Figure 9.1.4 - Mean Task Failures for 11-Digit Dialing Tasks by Lateral Pitch and Disability
(Groups A and B are significantly different overall.)

Stored Memory Recall Tasks

A two-way analysis revealed significant effects on task completion times due to disability ($F_{(4,29)} = 6.69$, $p < 0.007$), lateral pitch ($F_{(3,29)} = 6.07$, $p < 0.0005$), and the lateral pitch by disability interaction ($F_{(12,29)} = 2.06$, $p < 0.02$). One-way analysis, overall, revealed a significant difference in task completion times between the 12 mm and 13 mm levels of lateral pitch ($p < 0.008$), with overall task completion times the fastest when using the 12 mm lateral pitch. This trend was evident for all participant groups, except for the SU participant group. The SU participant group performed the fastest with the 11 mm lateral pitch, and the slowest with the 12 mm lateral pitch (Figure 9.1.5).

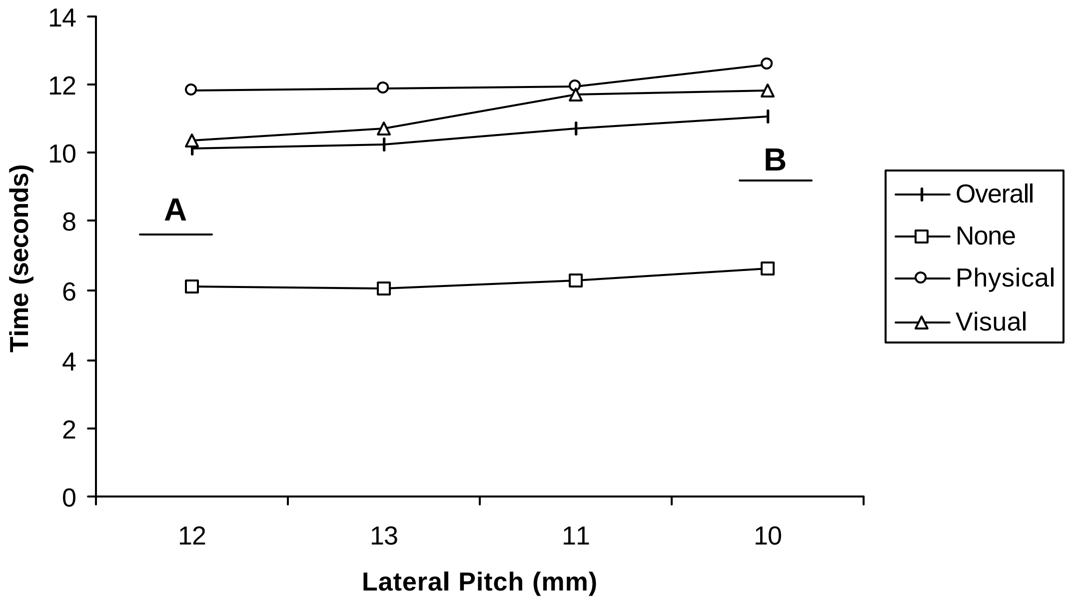


Figure 9.1.5 - Mean Stored Memory Recall Dialing Time by Lateral Pitch and Disability (Groups A and B are significantly different overall.)

A two-way analysis revealed significant effects on the number of task failures for lateral pitch ($F_{(3,29)} = 4.14$, $p < 0.006$), and the lateral pitch by disability interaction ($F_{(12,29)} = 8.03$, $p < 0.0001$), but not for disability. One-way analyses revealed a significant difference in the overall number of task failures between the 10 mm and 12 mm lateral pitch ($p < 0.017$), with the number of task failures the lowest when using the 12 mm lateral pitch. Participants in the Visual category committed the most task failures when using the 10 mm lateral pitch (Figure 9.1.6). Participants in the None and Physical categories did not commit any task failures when performing the stored memory recall tasks.

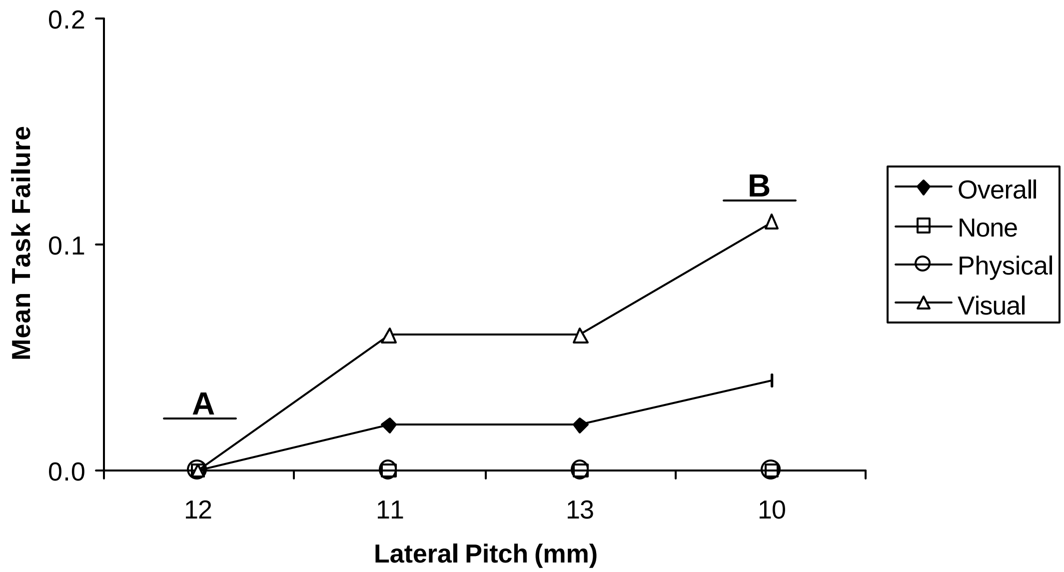


Figure 9.1.6 - Mean Task Failures for Stored Memory Recall Dialing Tasks by Lateral Pitch and Disability (Groups A and B are significantly different overall.)

911 Dialing Tasks

A two-way analysis revealed significant effects on 911 task completion time due to disability ($F_{(4,29)} = 6.20$, $p < 0.009$), lateral pitch ($F_{(3,29)} = 3.73$, $p < 0.01$), and the lateral pitch by disability interaction ($F_{(12,29)} = 4.80$, $p < 0.0001$). One-way analysis did not reveal any significant differences overall.

Overall task completion times were the fastest when using the 12 mm lateral pitch, followed closely by the 10 mm lateral pitch. Participants in the Physical category deviated from this trend slightly, and completed tasks the fastest when using the 11 mm and 13 mm levels of lateral pitch (Figure 9.1.7).

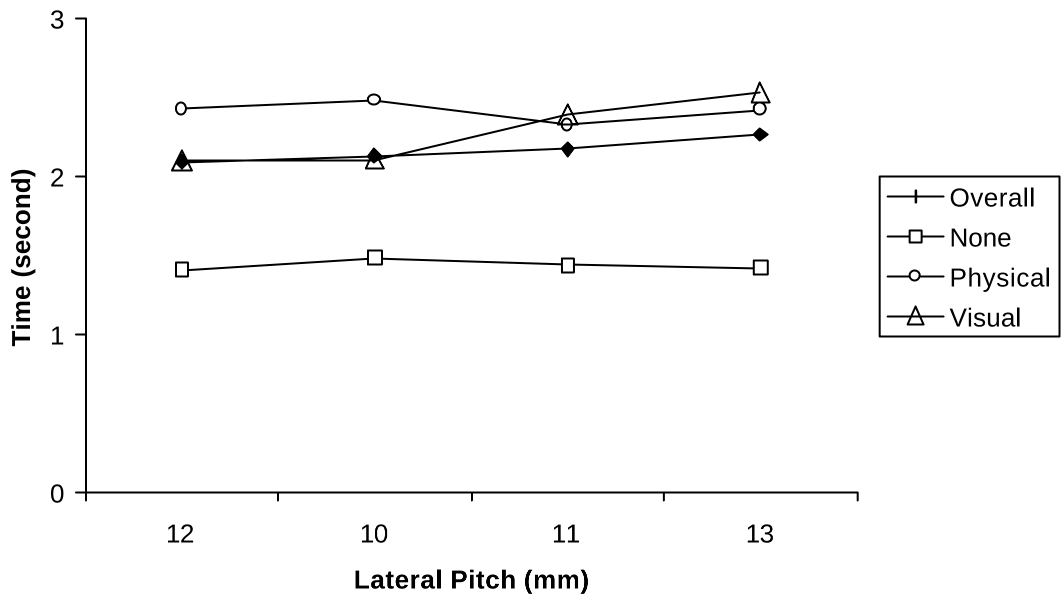


Figure 9.1.7 - Mean 911Dialing Time by Lateral Pitch and Disability

A two-way analysis revealed that the number of task failures was significantly affected by lateral pitch ($F_{(3,29)} = 7.76$, $p < 0.0001$), and the lateral pitch by disability interaction ($F_{(12,29)} = 6.74$, $p < 0.0001$), but not by disability. Overall one-way analyses revealed significant differences in the number of task failures between the 12 mm lateral pitch and all other levels ($p < 0.0002$).

The overall number of task failures was the lowest with the 10 mm lateral pitch, and highest with the 12 mm (Figure 9.1.8). This trend was evident for participants in the Visual category, but not for participants in the other categories. Those in the SU participant group had the highest number of task failures when using the 13 mm lateral pitch. Those in the ND and MU participant groups did not commit any task failures when performing the 911 dialing tasks.

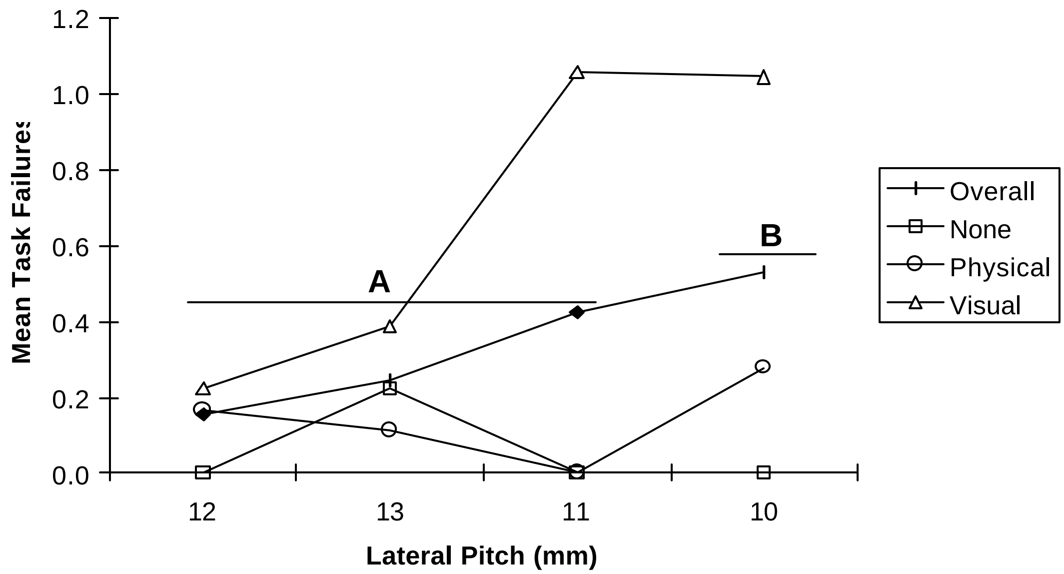


Figure 9.1.8 - Mean Task Failures for 911 Dialing Tasks by Lateral Pitch and Disability
(Groups A and B are significantly different overall.)

Total Time

Two-way analyses for total task completion time (the sum of the times to complete all tasks) revealed significant effected for disability ($F_{(4,29)} = 8.11$, $p < 0.004$), lateral pitch ($F_{(3,29)} = 27.37$, $p < 0.0001$), and the lateral pitch by disability interaction ($F_{(12,29)} = 2.43$, $p < 0.004$). One-way analyses, overall, revealed a significant ($p < 0.01$) difference in total task completion times for the 12 mm level of lateral pitch, when compared to the 10 mm and 11 mm levels (Figure 9.1.9).

Overall task completion times were the fastest when using the 12 mm lateral pitch, and the slowest with the 10 mm. This trend was evident for all participant groups, except for the SU participant group, which performed the fastest with the 13 mm lateral pitch (Figure 9.1.9).

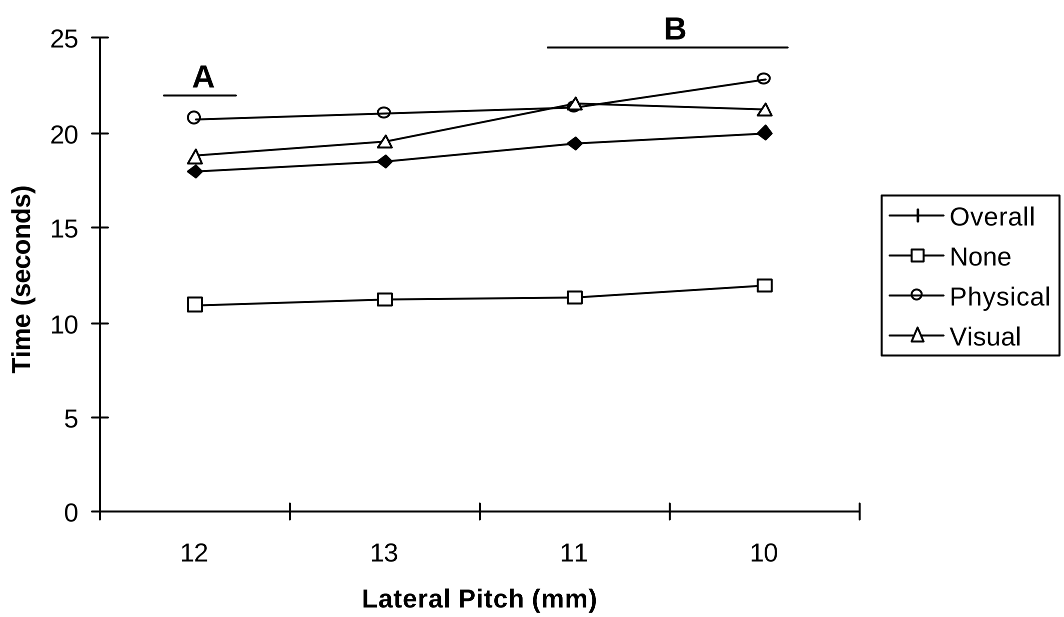


Figure 9.1.9 - Mean Time to Complete All Tasks by Lateral Pitch and Disability
(Groups A and B are significantly different overall.)

A two-way analysis revealed significant effects on the number of task failures due to disability ($F_{(4,29)} = 32.57$, $p < 0.0001$), lateral pitch ($F_{(3,29)} = 16.32$, $p < 0.0001$), and the lateral pitch by disability interaction ($F_{(12,29)} = 13.17$, $p < 0.0001$). One-way analyses, overall, revealed that the total number of task failures with the 12 mm and 13 mm levels of lateral pitch were significantly less than the total number of task failures with the 10 mm and 11 mm levels of lateral pitch ($p < 0.0001$).

The overall number of task failures was the lowest with the 12 mm lateral pitch (Figure 9.1.10). This trend held true for all participant groups, except for those in the SU group. Participants in the SU group had the lowest number of task failures when using the 11 mm and 13 mm levels of lateral pitch, and the highest when using the 10 mm and 12 mm levels. Participants in the MU group did not commit any task failures during the experiment.

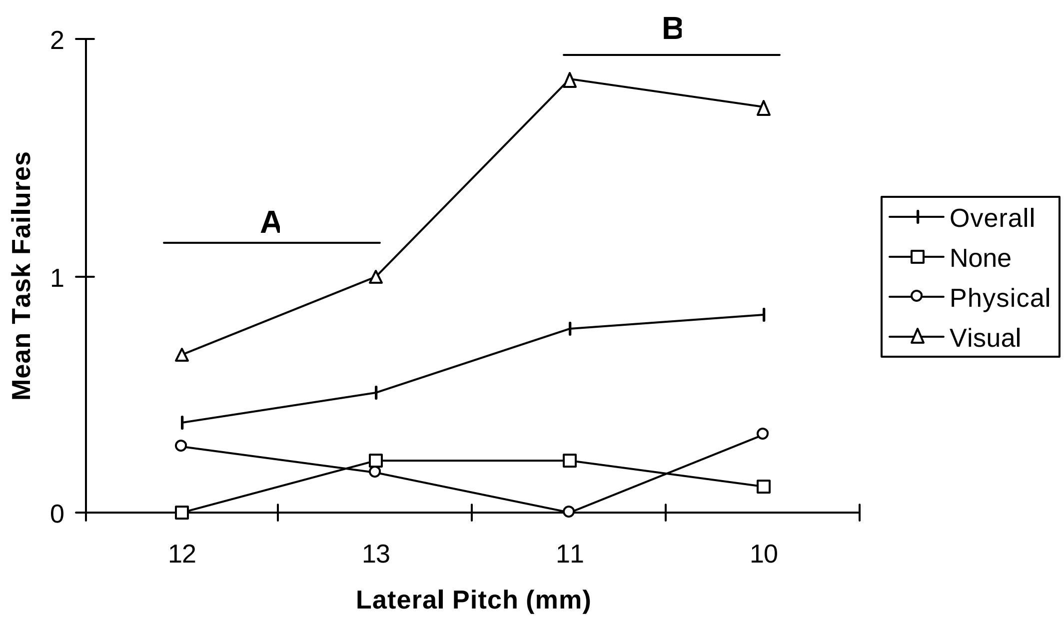


Figure 9.1.10 - Mean Total Task Failures for All Tasks by Lateral Pitch and Disability
(Groups A and B are significantly different overall.)

Preference Ranking

A two-way analysis revealed significant effects on preference for lateral pitch ($F_{(3,17)} = 24.55$, $p < 0.0001$), but not for disability or the lateral pitch by disability interaction. A one-way analysis, overall, revealed significant ($p < 0.0001$) differences between the 10 mm, 11 mm, and the 12 mm and 13 mm levels of lateral pitch (Groups A, B, and C, Figure 9.1.11). Across all participant groups, the 12 mm lateral pitch had the best preference ranking (lowest rank order), followed in order by the 13 mm, 11 mm, and 10 mm levels.

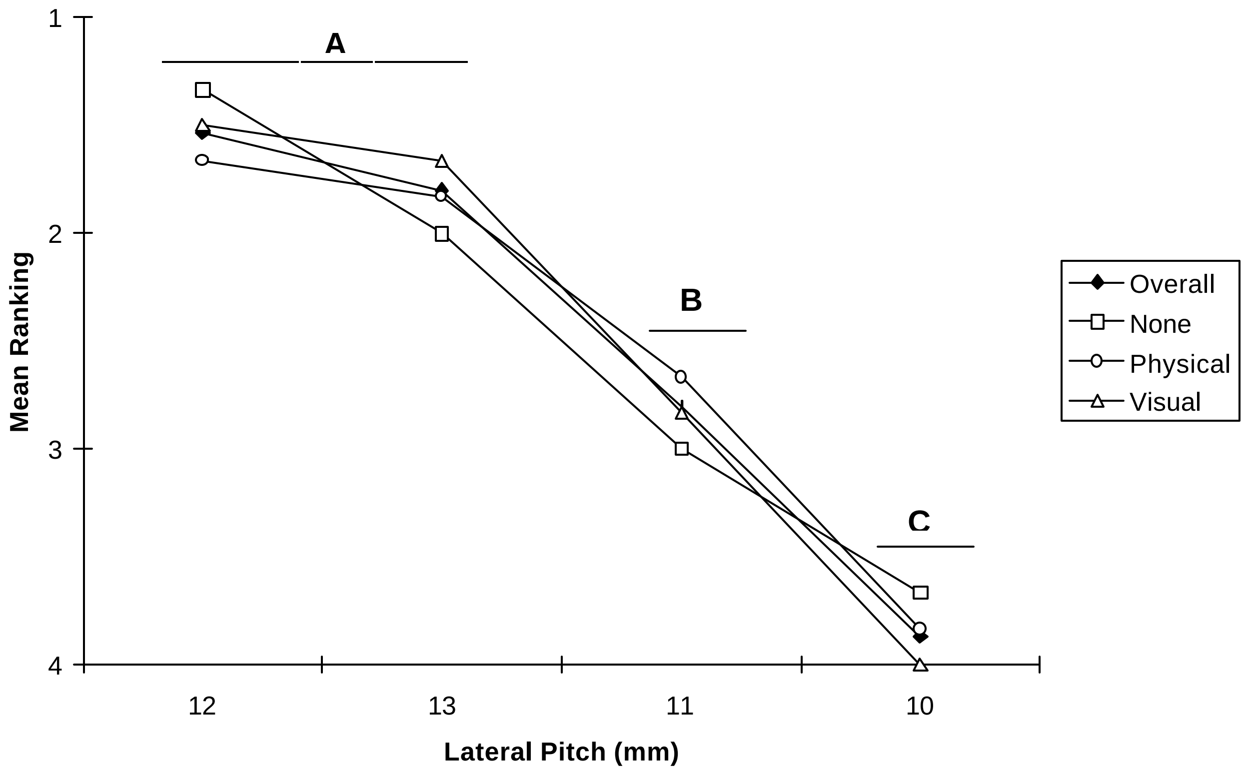


Figure 9.1.11 - Mean Preference Rankings by Lateral Pitch and Disability
(Lower numbers indicate greater preference. Groups A, B, and C are significantly different overall.)

Ease of Use Ratings

A two-way analysis revealed significant effects on ease of use ratings for disability ($F_{(4,17)} = 7.03$, $p < 0.006$), and lateral pitch ($F_{(3,17)} = 9.75$, $p < 0.0001$), but not the lateral pitch by disability interaction. An overall one-way analysis revealed a significant ($p < 0.0001$) difference in ease of use ratings between the 10 mm lateral pitch, and all other levels (Groups A and B, Figure 9.1.12). The 12 mm lateral pitch was rated the highest across all disability groups, followed in order by the 13 mm, 11 mm, and 10 mm levels.

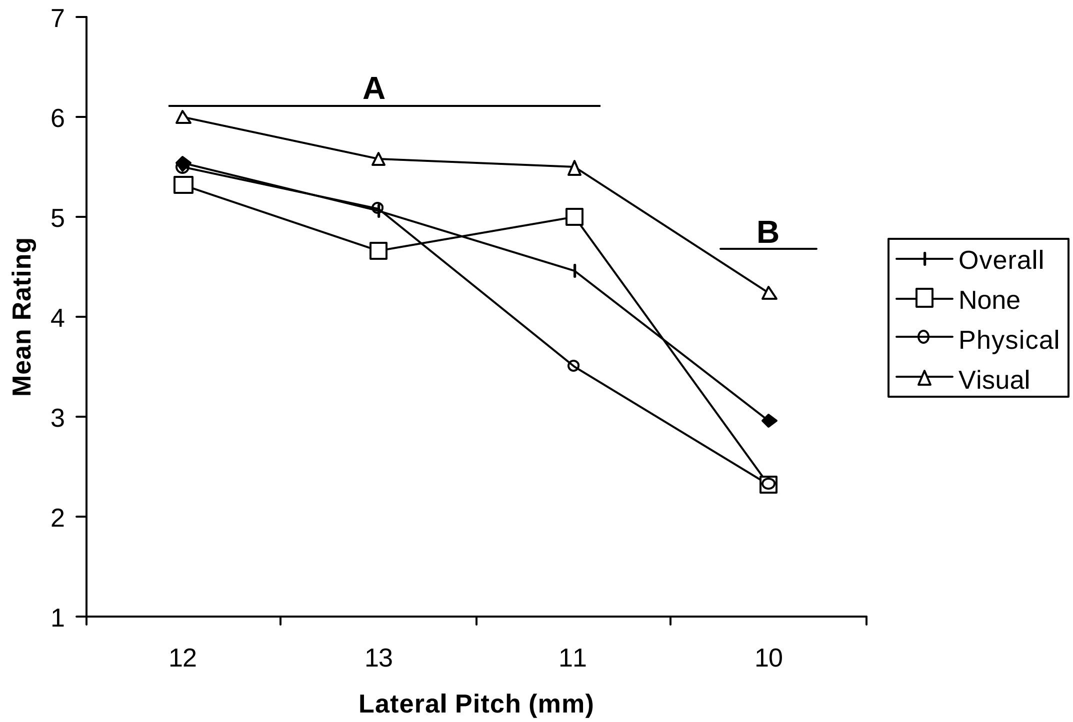


Figure 9.1.12 - Mean Ease of Use Ratings by Lateral Pitch and Disability
(Higher numbers indicate greater ease of use. Groups A and B are significantly different overall.)

Accessibility Ratings

A two-way analysis revealed significant effects on accessibility ratings for lateral pitch ($F_{(3,17)} = 13.64$, $p < 0.0001$), but not for disability ($F_{(4,17)} = 2.24$, $p < 0.14$) or the lateral pitch by disability interaction. A one-way analysis overall revealed a significant ($p < 0.0001$) difference in accessibility ratings between the 10 mm, 11 mm, and the 12 mm and 13 mm levels of lateral pitch (Groups A, B, and C, Figure 9.1.13). The 12 mm lateral pitch was rated the highest for accessibility overall, followed by the 13 mm, 11 mm and 10 mm levels.

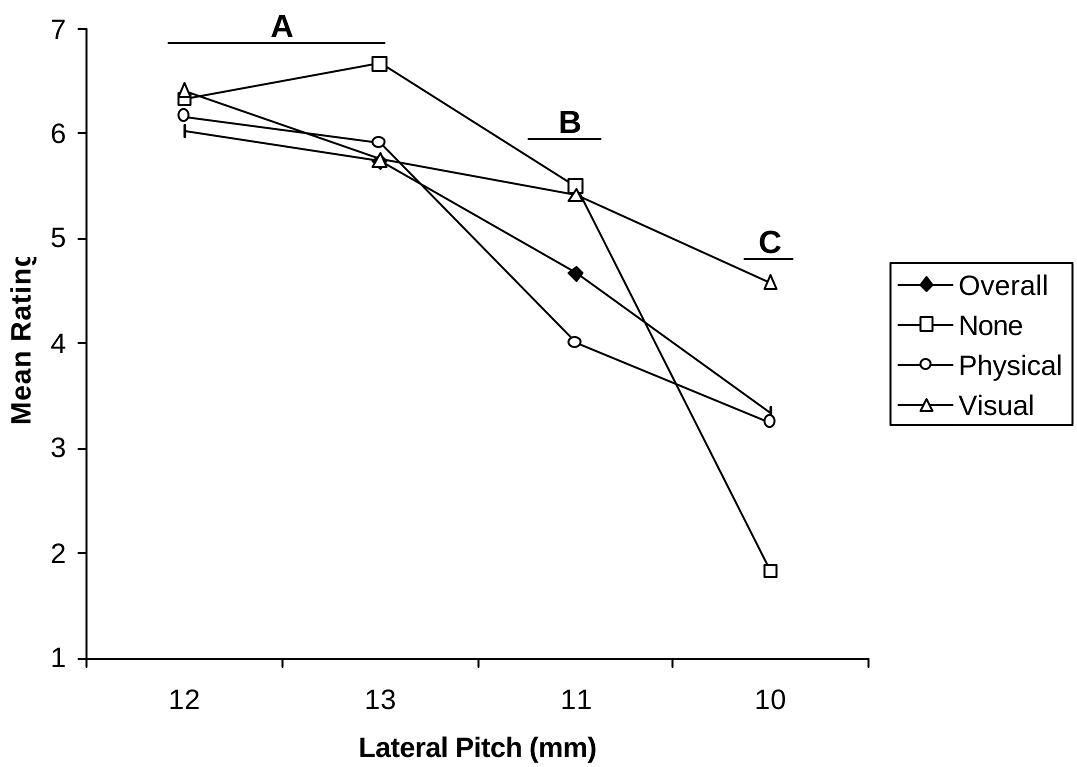


Figure 9.1.13 - Mean Accessibility Ratings by Lateral Pitch and Disability
(Higher numbers indicate greater accessibility. Groups A, B, and C are significantly different overall.)

9.2 - Effects of Keystroke

7-Digit Dialing Tasks

A two-way analysis of task completion times revealed significant effects for disability ($F_{(4,24)} = 5.85$, $p < 0.01$), but not for the keystroke or the keystroke by disability interaction. One-way analyses did not reveal any significant differences in task completion times due to keystroke overall, or for any participant group. Overall task completion times for each level of keystroke were nearly identical (Figure 9.2.1).

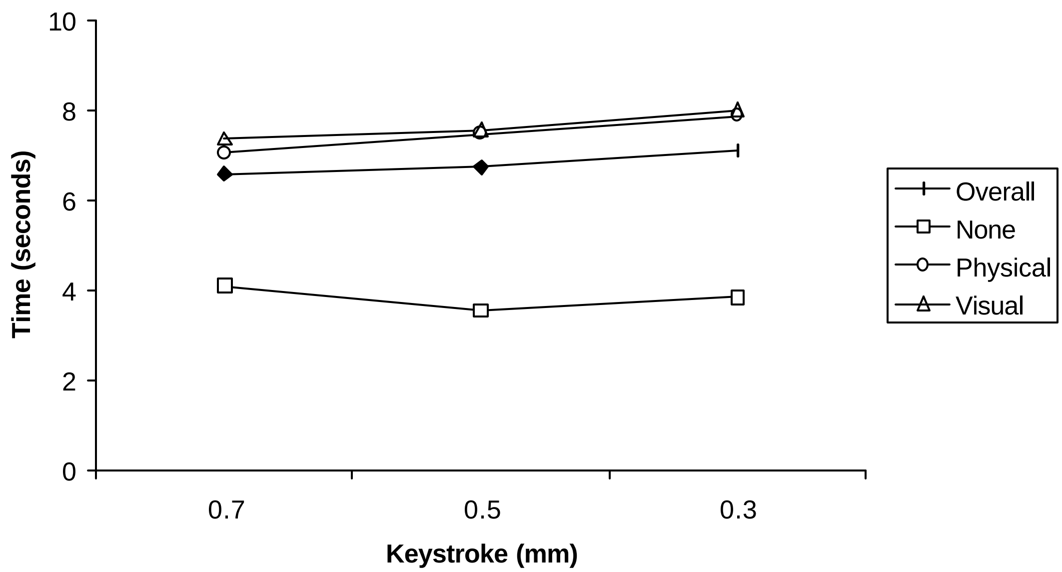


Figure 9.2.1 - Mean 7-Digit Dialing Time by Keystroke and Disability

A two-way analysis of the number of task failures revealed no significant effects due to disability, however the keystroke ($F_{(2,24)} = 5.00$, $p < 0.008$) and keystroke by disability interaction effects were significant ($F_{(8,24)} = 1.88$, $p < 0.07$). A one-way analysis, overall, revealed a significant ($p < 0.0001$) difference in the number of task failures between the 0.5 mm keystroke, and the 0.3 mm and 0.7 mm levels of keystroke (Groups A and B, Figure 9.2.2).

The number of task failures for those in the LB and SU participant groups were the lowest when using the 0.5 mm keystroke. The number of task failures for the remaining participant groups were similar for each level of keystroke (Figure 9.2.2).

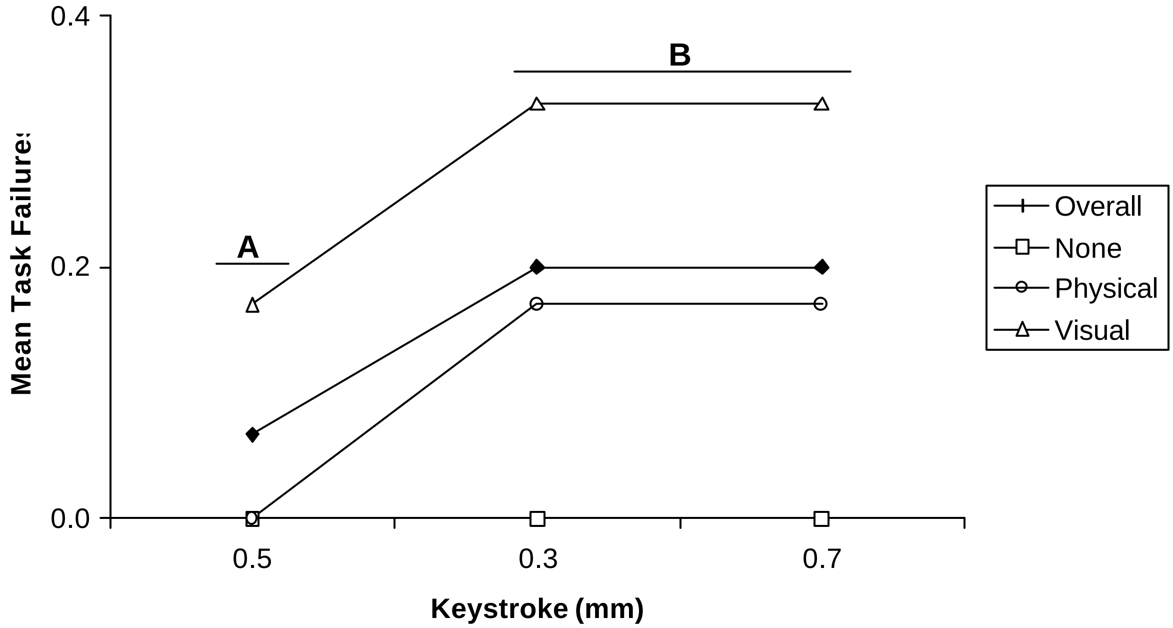


Figure 9.2.2 - Mean Task Failures for 7-Digit Dialing Tasks by Keystroke and Disability
(Groups A and B are significantly different overall.)

11-Digit Dialing Tasks

A two-way analysis of task completion times revealed significant effects due to disability ($F_{(4,24)} = 6.04$, $p < 0.01$), keystroke ($F_{(2,24)} = 3.73$, $p < 0.03$), and the keystroke by disability interaction ($F_{(8,24)} = 6.33$, $p < 0.0001$). One-way analysis, overall, did not reveal any significant differences in task completion times.

Overall task completion times were the fastest when using the 0.7 mm keystroke. Participants in the Visual category, however, exhibited the opposite trend and completed tasks more slowly when using the 0.7 mm keystroke (Figure 9.2.3).

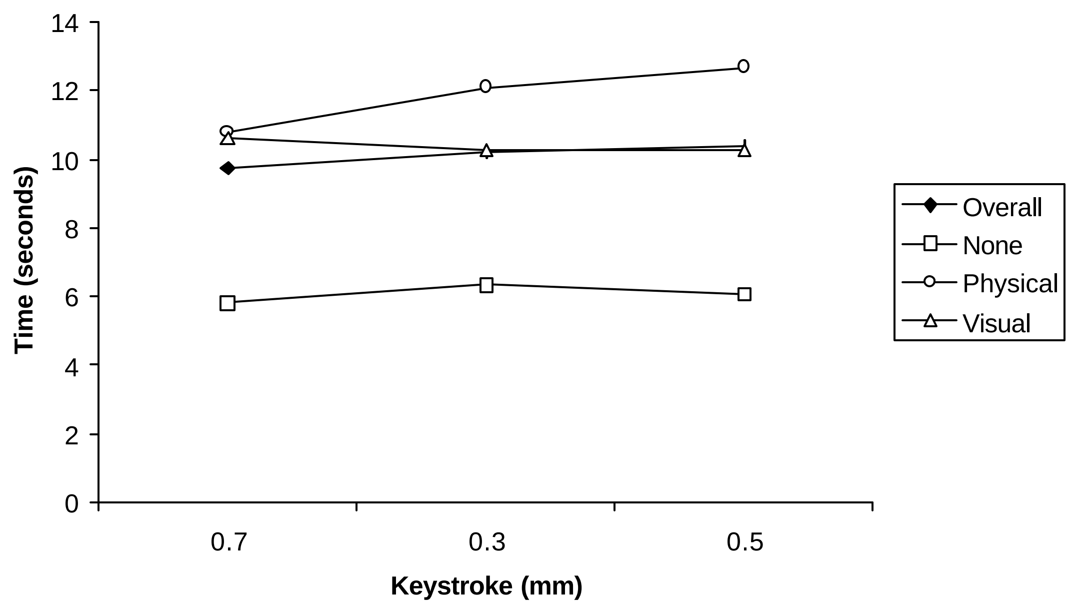


Figure 9.2.3 - Mean 11-Digit Dialing Time by Keystroke and Disability

A two-way analysis of the number of task failures revealed significant effects due to keystroke ($F_{(2,24)} = 47.5$, $p < 0.0001$) and the keystroke by disability interaction ($F_{(8,24)} = 28.75$, $p < 0.0001$), but not by disability. A one-way analysis, overall, revealed a significant ($p < 0.0001$) difference in the number of task failures between the 0.7 mm level of keystroke, when compared to the 0.3 mm and 0.5 mm levels (Groups A and B, Figure 9.2.4).

The overall number of task failures was the lowest when using the 0.5 mm keystroke, and the highest when using the 0.7 mm keystroke. Reported task failures reflect the performance of only those in the TB and SU participant groups (Figure 9.2.4), since participants in the remaining groups did not commit any task failures during this portion of the study.

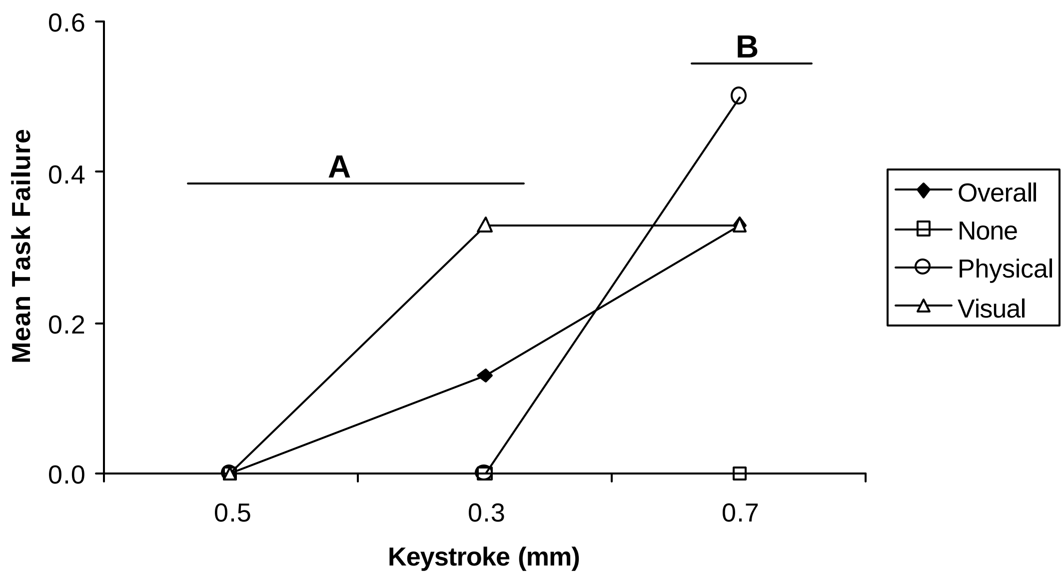


Figure 9.2.4 - Mean Task Failures for 11-Digit Dialing Tasks by Keystroke and Disability (Groups A and B are significantly different overall.)

Stored Memory Recall Tasks

A two-way analysis revealed significant effects on task completion times due to disability ($F_{(4,24)}$ = 3.63, p < 0.05), and keystroke ($F_{(2,24)}$ = 2.39, p < 0.09), however the keystroke by disability interaction was not significant. One-way analyses did not reveal any significant differences due to keystroke. All participant groups performed the fastest when using the 0.3 mm keystroke, however differences in task completion time between each level of keystroke were minor (Figure 9.2.4).

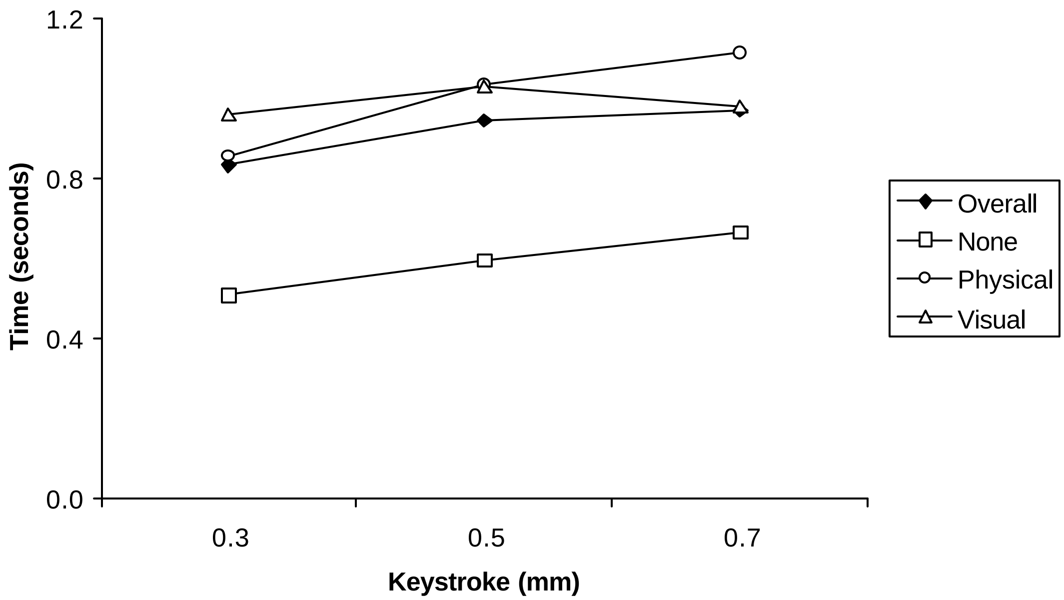


Figure 9.2.4 - Mean Stored Memory Recall Dialing Time by Keystroke and Disability

A statistical analysis of the number or task failures was not performed. Participants did not commit any task failures when performing stored memory recall tasks.

911 Dialing Tasks

A two-way analysis of task completion times revealed significant effects for disability ($F_{(4,24)} = 3.95$, $p < 0.04$), but not for keystroke or the keystroke by disability interaction. One-way analyses did not reveal any significant differences between levels of keystroke.

Participants in the TB and SU groups completed 911 dialing tasks more slowly than those in the other participant groups. Task completion times were the fastest with the 0.7 mm keystroke overall, however differences in task completion times due to keystroke were minor (Figure 9.2.5).

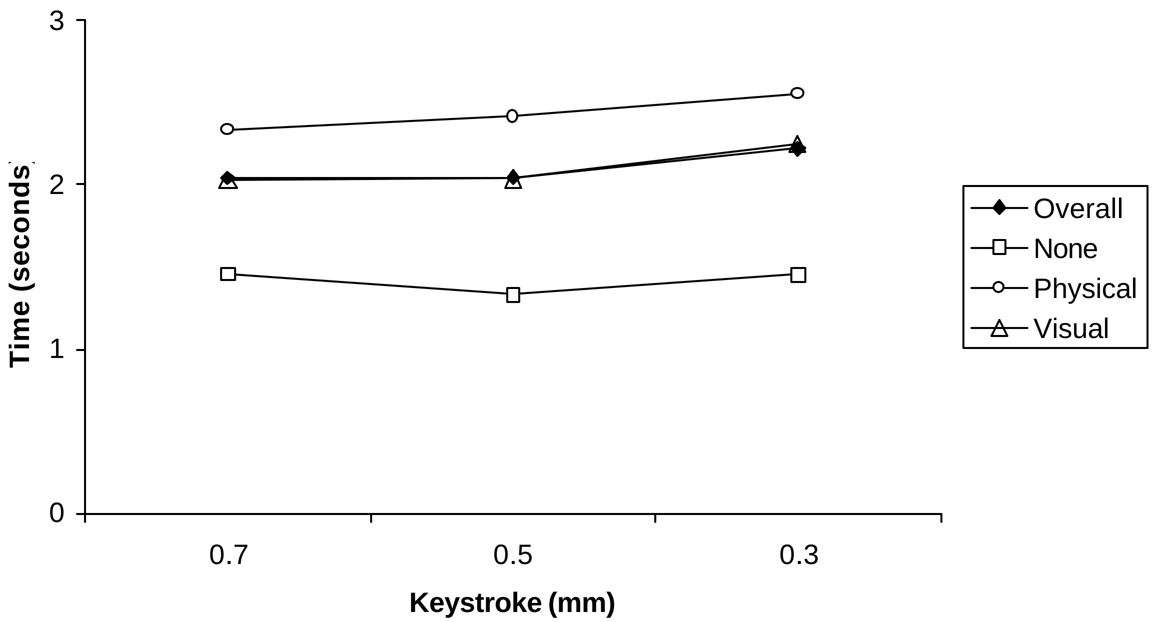


Figure 9.2.5 - Mean 911 Dialing Time by Keystroke and Disability

A two-way analysis revealed that the number of task failures was significantly affected by keystroke ($F_{(2,24)} = 8.57$, $p < 0.0003$) and the keystroke by disability interaction ($F_{(8,24)} = 5.00$, $p < 0.0001$), but not by disability. One-way analysis, overall, revealed a significant ($p < 0.004$) difference in the number of task failures between the 0.3 mm and 0.7 mm levels of keystroke (Groups A and B, Figure 9.2.6).

The overall number of task failures was the lowest when using the 0.3 mm keystroke, and the highest when using the 0.7 mm keystroke. The number of task failures reflects the performance of those in the Visual category, since participants in the other categories did not commit any task failures.

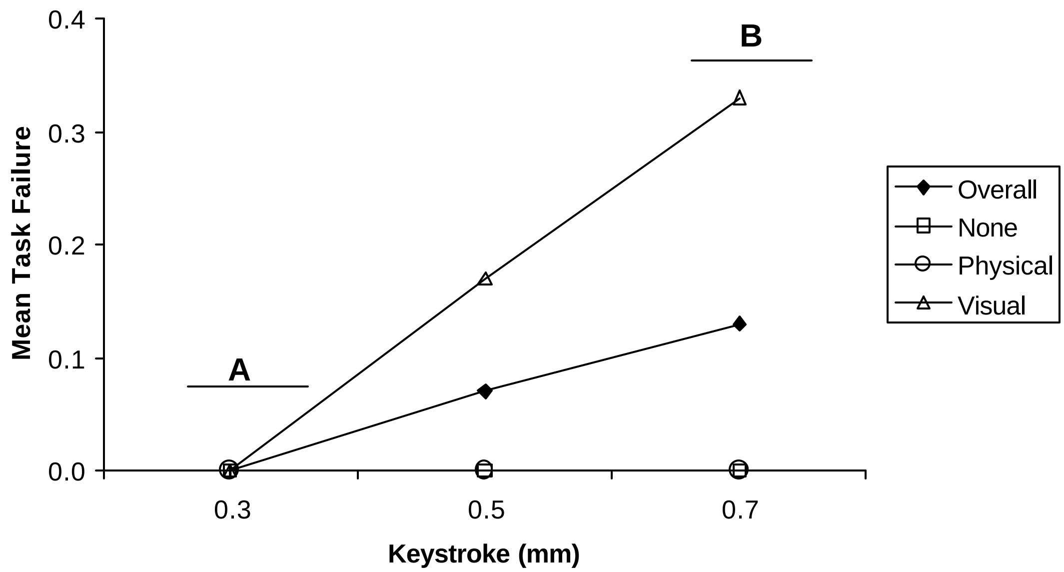


Figure 9.2.6 - Mean Task Failures for 911 Dialing Tasks by Keystroke and Disability
(Groups A and B are significantly different overall.)

Total Task Completion Time

A two-way analysis revealed significant effects for total task completion time for disability ($F_{(4,24)} = 6.39$, $p < 0.008$), and the keystroke by disability interaction ($F_{(8,24)} = 2.85$, $p < 0.005$) but not for keystroke. One-way analyses did not reveal any significant differences between each level of keystroke, however. Overall task completion times were the fastest with the 0.7 mm keystroke, however differences in task completion times between each level of keystroke were minor (Figure 9.2.7).

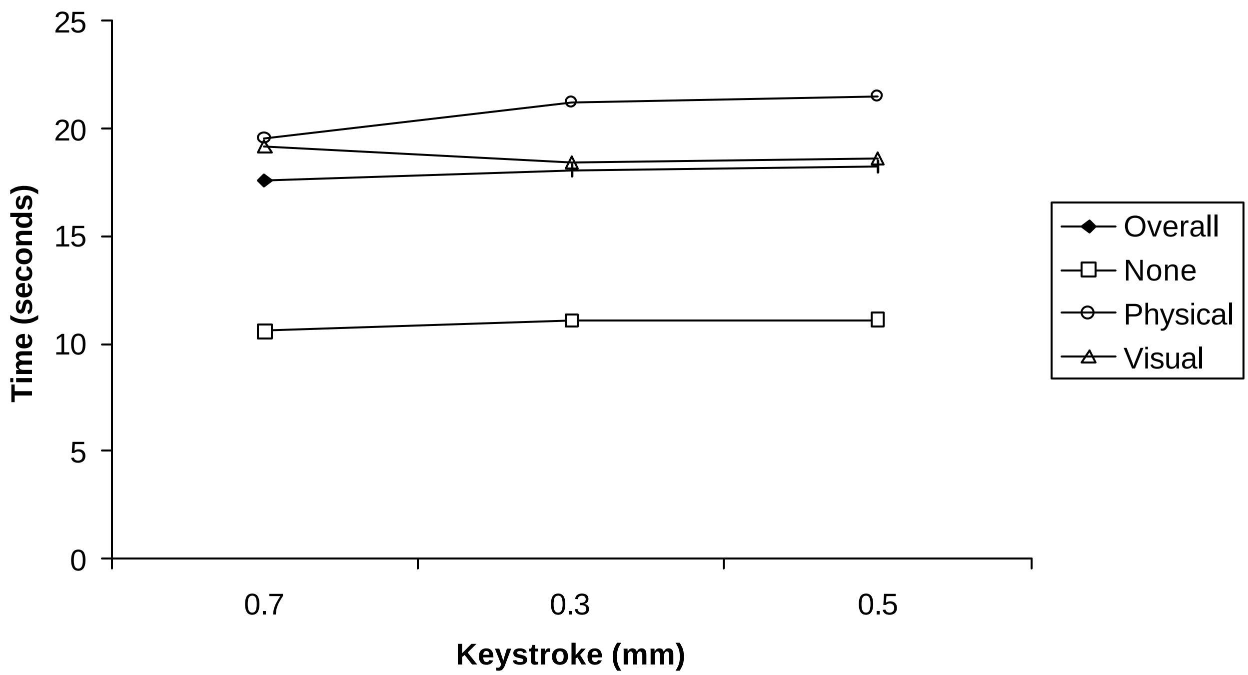


Figure 9.2.7 - Mean Time to Complete All Tasks by Keystroke and Disability

A two-way analysis of the number of task failures revealed significant effects for disability ($F_{(4,24)} = 2.86$, $p < 0.08$), keystroke ($F_{(2,24)} = 33.79$, $p < 0.0001$), and the keystroke by disability interaction ($F_{(8,24)} = 7.93$, $p < 0.0001$). A one-way analysis, overall, revealed a significant ($p < 0.0001$) difference in the total number of task failures between the 0.5 mm and 0.7 mm levels of keystroke (Groups A and B, Figure 9.2.8). The overall number of task failures was the lowest when using the 0.5 mm keystroke, and highest when using the 0.7 mm keystroke.

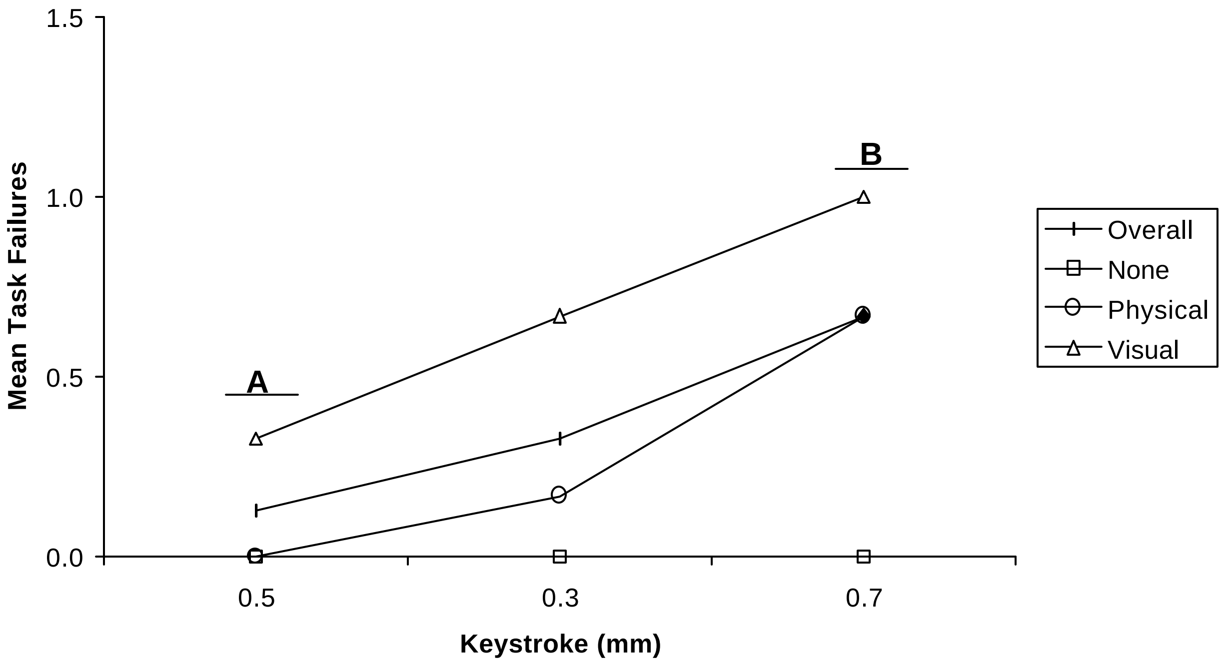


Figure 9.2.8 - Mean Task Failures for All Tasks by Keystroke and Disability
(Groups A and B are significantly different overall.)

Preference Rankings

Preference rankings were not significantly affected by disability, keystroke, or the keystroke by disability interaction. Participants preferred the 0.5 mm keystroke most often, followed by the 0.7 mm and 0.3 mm levels. Differences in keystroke ranking overall, and by disability group were not significant (Figure 9.2.9).

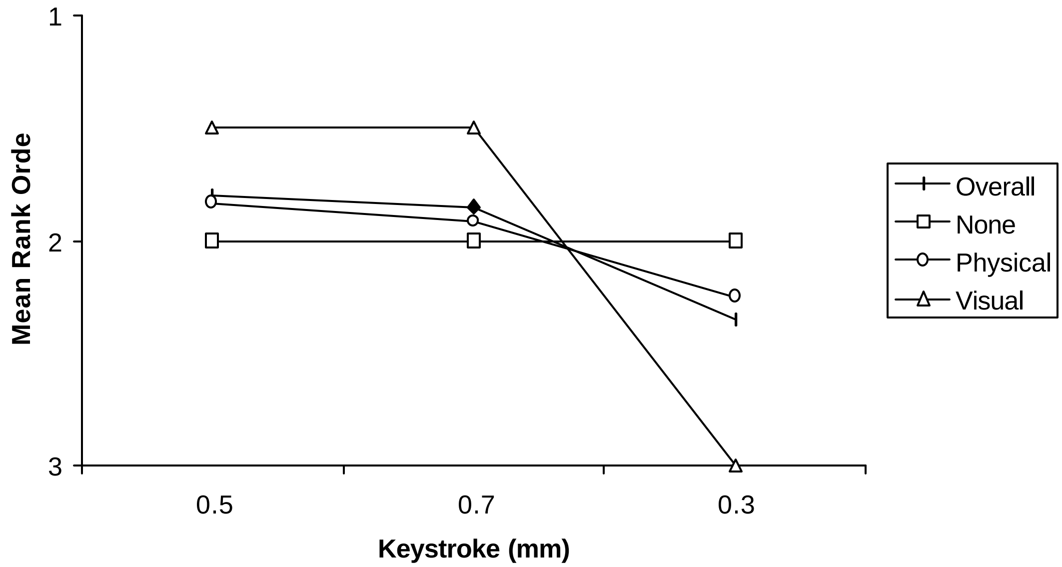


Figure 9.2.9 - Mean Preference Rankings by Keystroke and Disability

Ease of Use Ratings

Ease of use ratings were not significantly affected by disability, keystroke, or the keystroke by disability interaction. Most participant groups rated the 0.7 mm keystroke the highest for ease of use. Participants in the Physical category, however, rated the 0.5 mm keystroke the highest for ease of use. Differences in ease of use ratings overall, and by disability group were not significant (Figure 9.2.10).

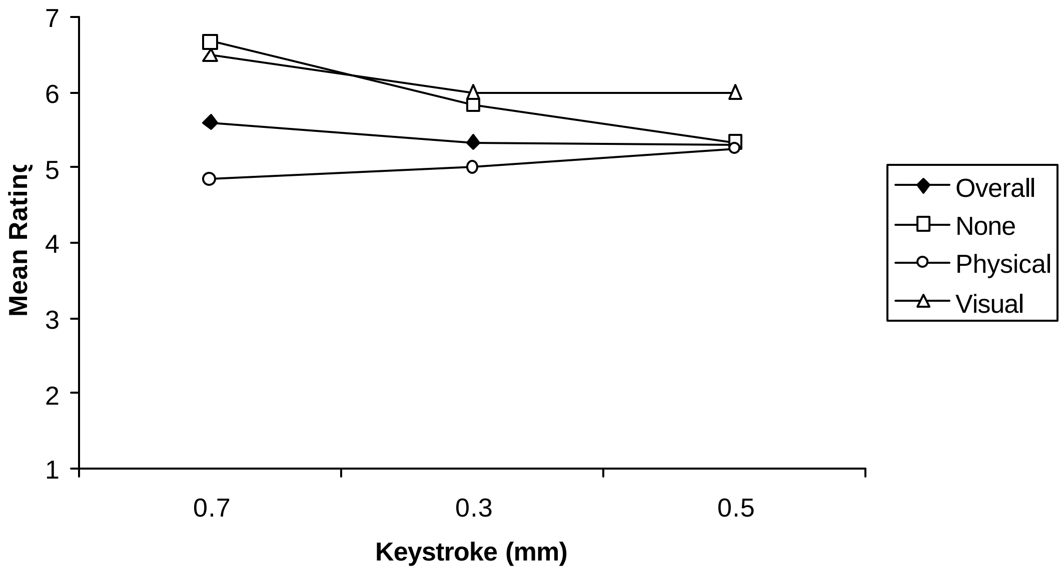


Figure 9.2.10 - Mean Ease of Use Ratings by Keystroke and Disability

Accessibility Ratings

Accessibility ratings were not significantly affected by disability, keystroke, or the keystroke by disability interaction. Participants generally rated the 0.7 mm keystroke the highest for accessibility, although those in the Physical category rated the 0.3 mm keystroke the highest. Differences in accessibility ratings overall, and by disability group were not significant (Figure 9.2.11).

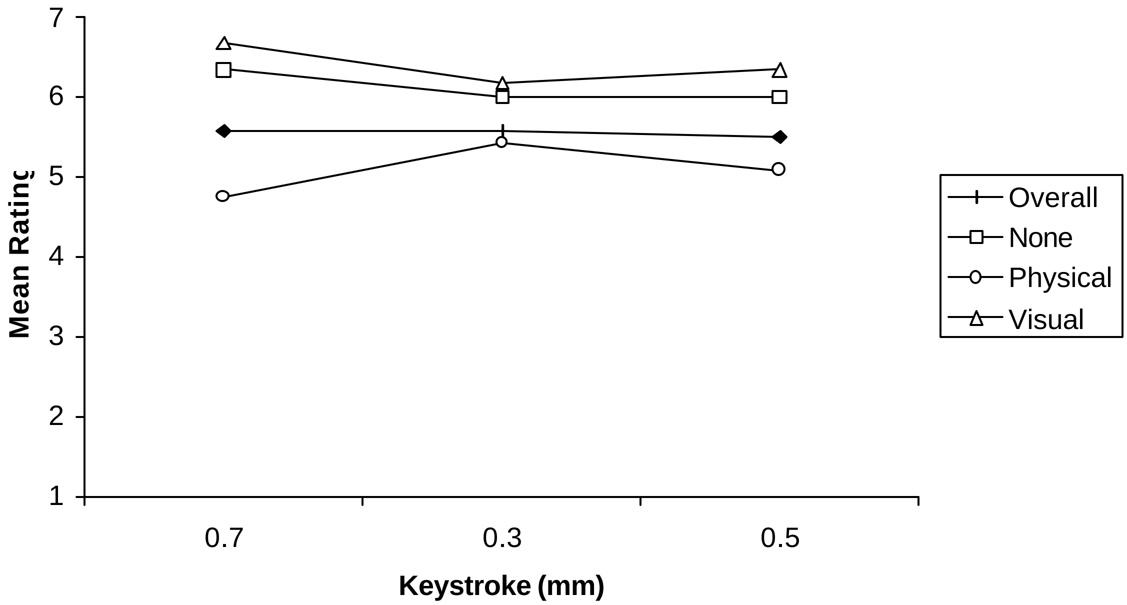


Figure 9.2.11 - Mean Accessibility Ratings by Keystroke and Disability

9.3 - Effects of Key Height

7-Digit Dialing Tasks

A two-way analysis of task completion times revealed significant effects for disability ($F_{(4,24)}$ = 10.25, $p < 0.002$), and the key height by disability interaction ($F_{(8,24)}$ = 3.28, $p < 0.002$), but not for key height. One-way analyses, overall, did not reveal any significant differences in task completion time due to key height. Participants completed dialing tasks the fastest overall when using the 0.7 mm key height, and the slowest when using the 0.3 mm key height (Figure 9.3.1).

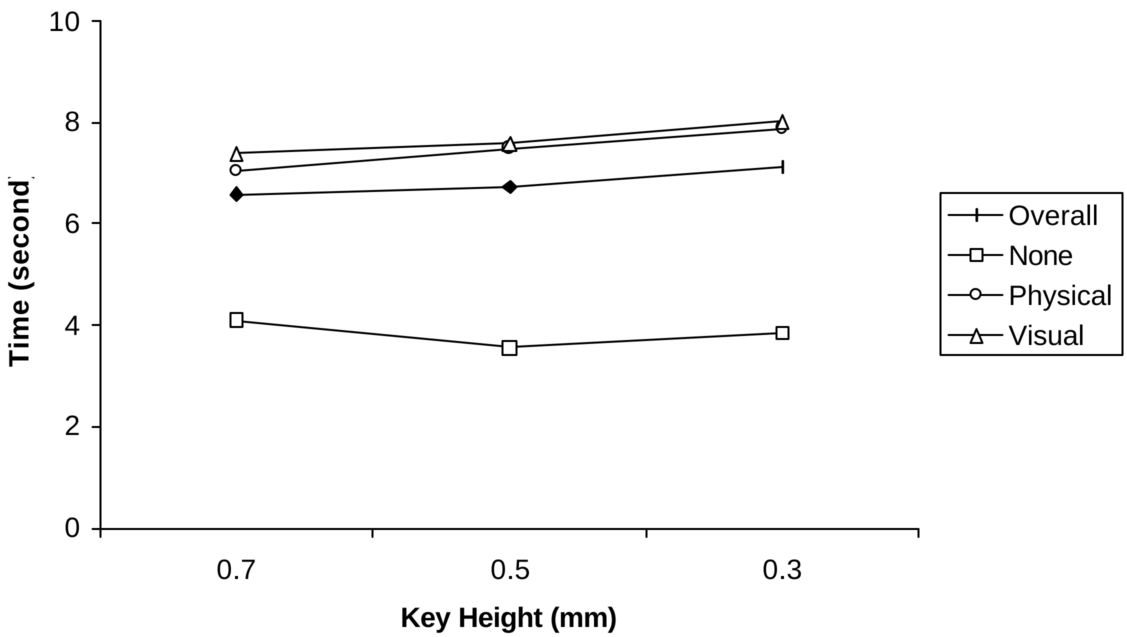


Figure 9.3.1 - Mean 7-Digit Dialing Time by Key Height and Disability

A two-way analysis of the number of task failures revealed significant effects for disability ($F_{(4,24)} = 4.28$, $p < 0.03$), key height ($F_{(2,24)} = 14.74$, $p < 0.0001$), and the key height by disability interaction ($F_{(8,24)} = 6.84$, $p < 0.0001$). One-way analysis, overall, revealed significant ($p < 0.001$) differences in the number of task failures between the 0.7 mm and 0.3 mm key height, when compared to the 0.5 mm key height (Groups A and B, Figure 9.3.2).

The number of task failures was the lowest when using the 0.7 mm key height, and the highest when using the 0.5 mm key height. This trend was evident for participants in the Visual category, and overall. Participants in the None category, however, had the lowest number of task failures when using the 0.3 mm key height. Participants in the Physical category did not commit any task failures with the different levels of key height.

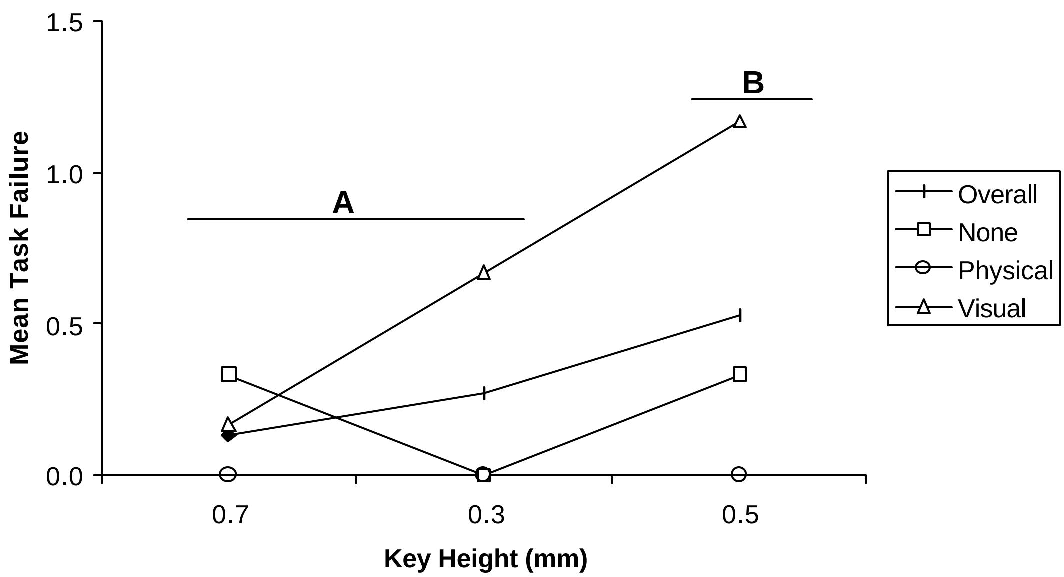


Figure 9.3.2 - Mean Task Failures for 7-Digit Dialing Tasks by Key Height and Disability
(Groups A and B are significantly different overall.)

11-Digit Dialing Tasks

A two-way analysis of task completion times revealed significant effects for disability ($F_{(4,24)} = 10.13$, $p < 0.002$), and key height ($F_{(2,24)} = 4.51$, $p < 0.01$), but not for the key height by disability interaction. One-way analyses did not reveal any significant differences overall, or by disability category. Participants generally completed tasks the fastest when using the 0.7 mm key height, and the slowest when using the 0.3 mm height (Figure 9.3.3).

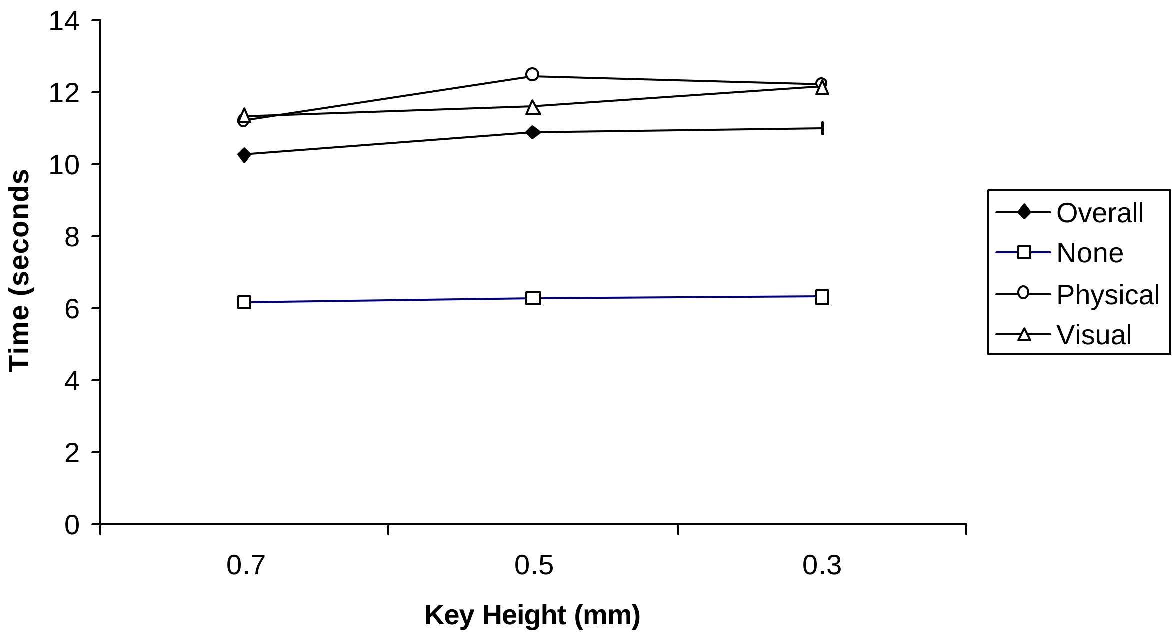


Figure 9.3.3 - Mean 11-Digit Dialing Time by Key Height and Disability

A two-way analysis of the number of task failures revealed significant effects for disability ($F_{(4,24)} = 13.19$, $p < 0.0005$), and the key height by disability interaction ($F_{(8,24)} = 3.63$, $p < 0.0006$), but not for key height. One-way analysis, overall, did not reveal any significant differences in the number of task failures, however.

The number of task failures for those in the Visual category was the lowest when using the 0.3 mm key height, and the highest when using the 0.7 mm key height (Figure 9.3.4). Participants in the other disability categories did not commit any task failures during this portion of the study.

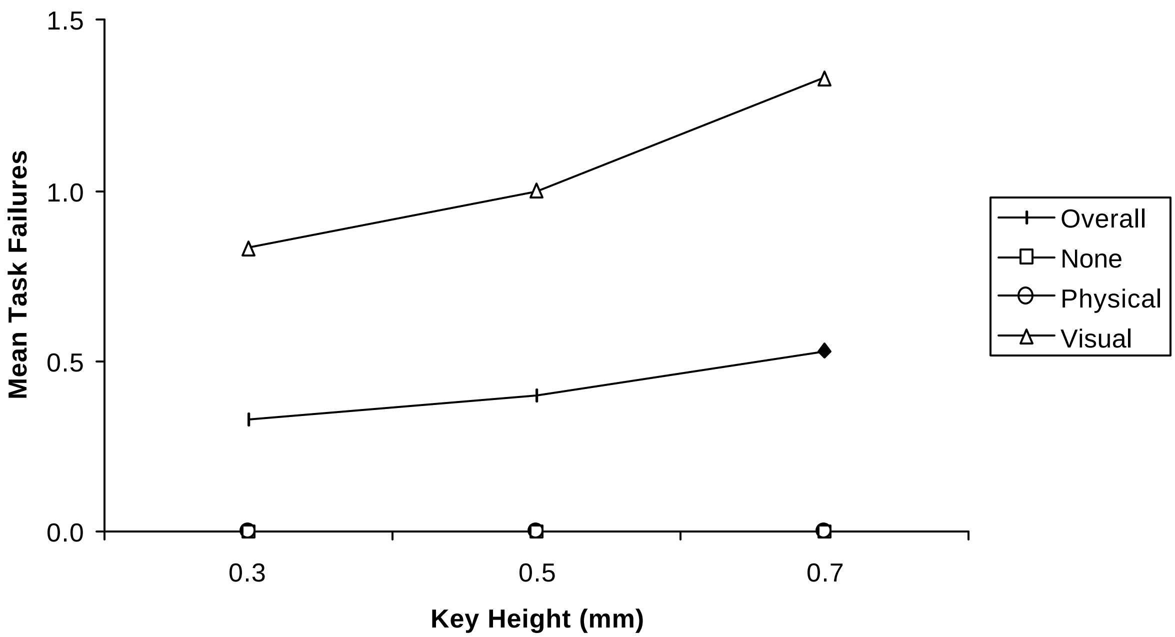


Figure 9.3.4 - Mean Task Failures for 11-Digit Dialing Tasks by Key Height and Disability

Stored Memory Recall Tasks

A two-way analysis of task completion times revealed significant effects for disability ($F_{(4,24)} = 8.17$, $p < 0.003$), and key height ($F_{(2,24)} = 3.99$, $p < 0.02$), but not for the key height by disability interaction ($F_{(8,24)} = 1.58$, $p < 0.14$). One-way analysis, overall, did not reveal any significant differences overall due to key height, however.

Overall task completion times were the fastest when using the 0.7 mm key height. The 0.5 mm key height produced the fastest task completion times for participants in the Visual category, and the slowest for those in the SU participant group. Participants in the TB group completed tasks significantly faster with the 0.5 mm key height, when compared to the 0.3 mm key height (Figure 9.3.5).

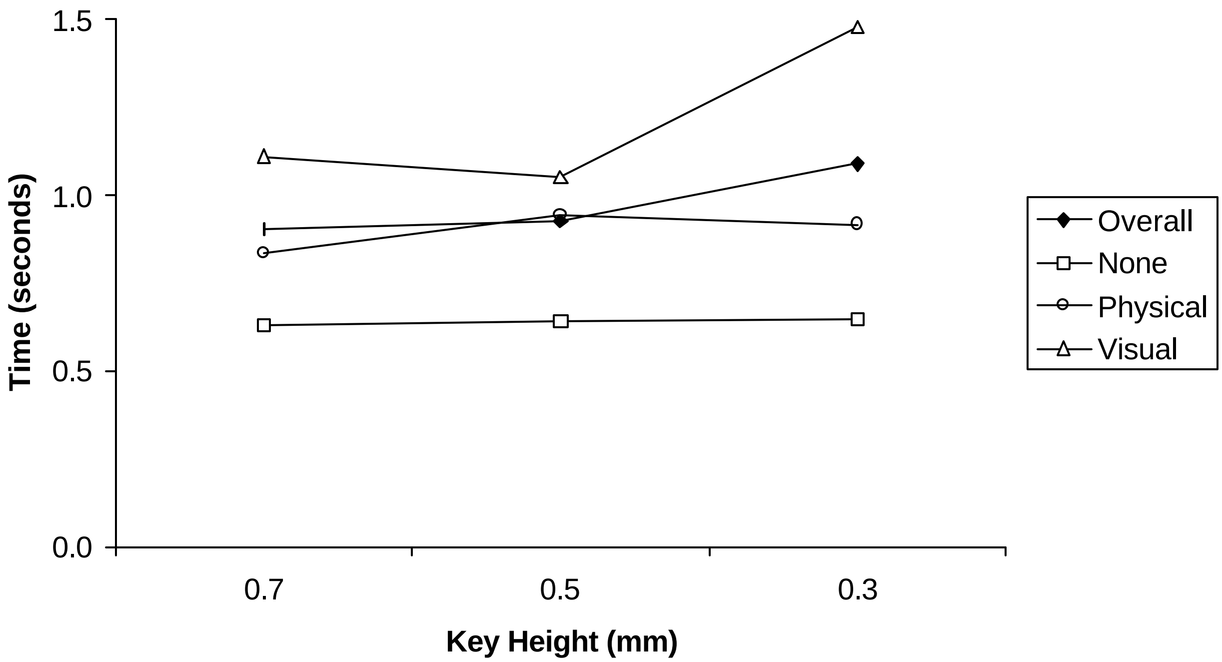


Figure 9.3.5 - Mean Stored Memory Recall Dialing Time by Key Height and Disability

A two-way analysis of the number of task failures revealed significant effects for key height ($F_{(2,24)}$ = 10.00 $p < 0.0001$), and the key height by disability interaction ($F_{(8,24)}$ = 10.00, $p < 0.0001$), but not for disability. A subsequent one-way analysis, overall, revealed a significant difference in the number of task failures ($p < 0.006$), however only participants in the LB group committed task failures during these tasks (Figure 9.3.6).

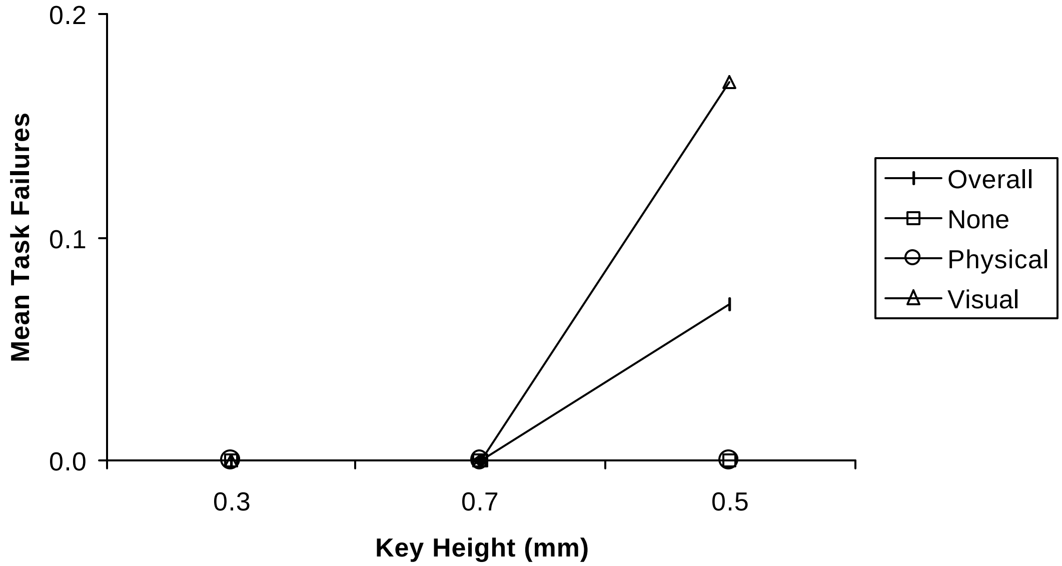


Figure 9.3.6 - Mean Task Failures for Stored Memory Recall Tasks by Key Height and Disability

911 Dialing Tasks

A two-way analysis of task completion times revealed significant effects for disability ($F_{(4,24)} = 4.57$, $p < 0.02$), however the key height and the key height by disability interaction were not significant. An overall one-way analysis did not reveal any significant differences due to key height, however.

Overall task completion times were the fastest when using the 0.7 mm key height. Participants in the TB participant group, however, had the fastest task completion times with the 0.5 mm key height (Figure 9.3.7).

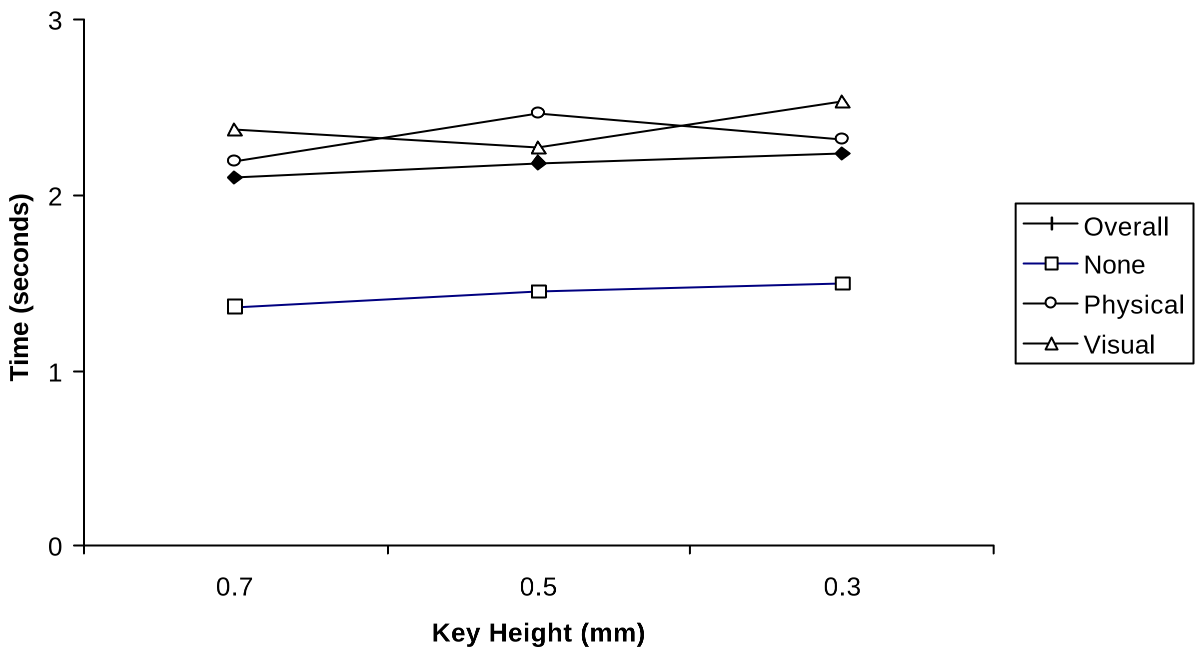


Figure 9.3.7 - Mean 911 Dialing Time by Key Height and Disability

A two-way analysis of the number of task failures revealed significant effects for key height ($F_{(2,24)}$ = 10.00 p < 0.0001), and the key height by disability interaction ($F_{(8,24)}$ = 10.00, p < 0.0001), but not for disability. One-way analysis, overall, did not reveal any significant differences in the number of task failures due key height, however. Task failures only occurred when participants in the TB group used the 0.5 mm key height (Figure 9.3.8).

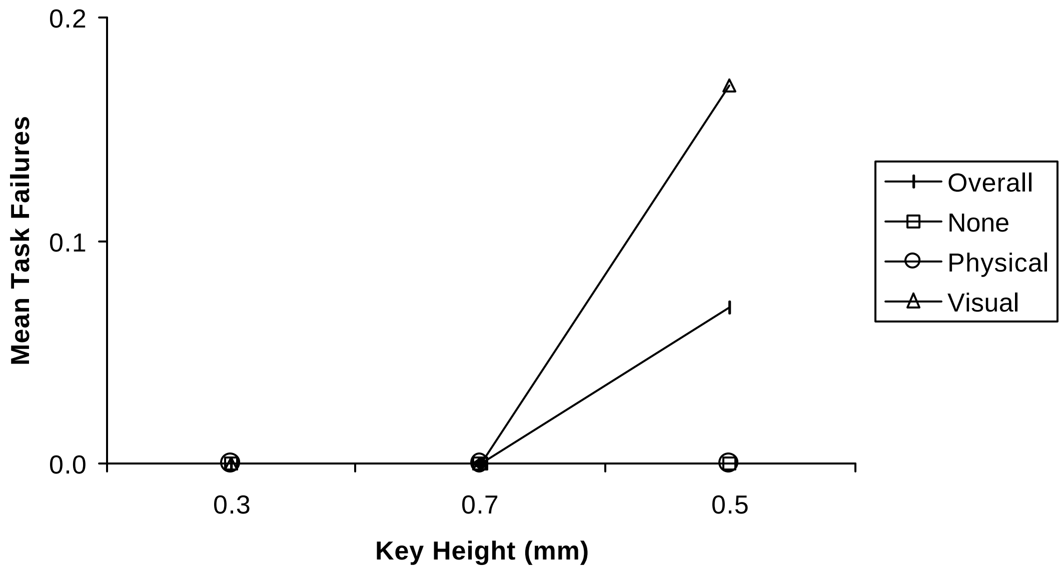


Figure 9.3.8 - Mean Task Failures for 911 Dialing Tasks by Key Height and Disability

Total Task Completion Time

A two-way analysis of total task completion times revealed significant effects for disability ($F_{(4,24)} = 10.80$, $p < 0.001$), and key height ($F_{(2,24)} = 3.96$, $p < 0.02$), but not for the key height by disability interaction ($F_{(8,24)} = 1.67$, $p < 0.11$). A subsequent one-way analysis, overall, did not reveal a significant difference in levels of key height, however.

Participants in the LB and ND groups completed tasks the fastest when using the 0.5 mm key height. Participants in the SU, MU, and TB groups completed tasks the fastest when using the 0.7 mm key height. The 0.7 mm key height provided the fastest task completion times, however the difference in completion times between the 0.7 mm level and the other levels was not significant (Figure 9.3.9).

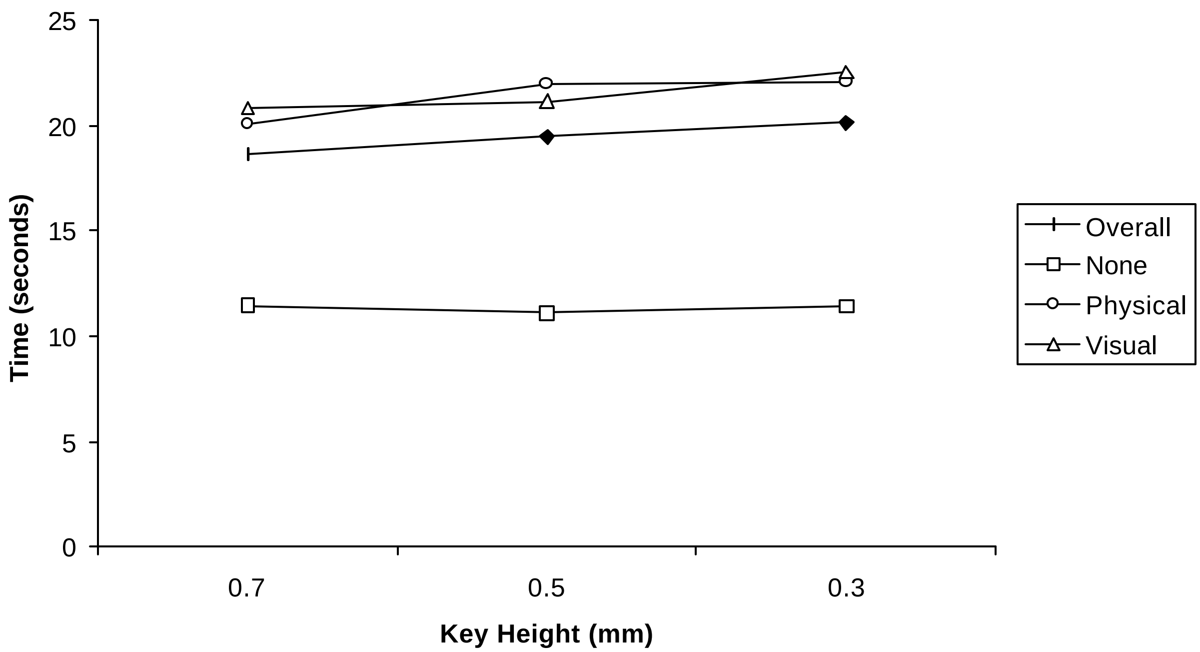


Figure 9.3.9 - Mean Time to Complete All Tasks by Key Height and Disability

A two-way analysis of the number of task failures revealed significant effects for disability ($F_{(4,24)} = 26.67$ $p < 0.0001$) key height ($F_{(2,24)} = 5.41$ $p < 0.005$), and the key height by disability interaction ($F_{(8,24)} = 2.89$, $p < 0.005$). Subsequent one-way analysis, however, did not reveal a significant difference in the overall number of task failures due to key height ($p < 0.11$).

The overall number of task failures was the lowest when using the 0.3 mm key height. The number of task failures for those in the LB participant group was the lowest when using the 0.7 mm key height, and the highest when using the 0.5 mm key height. Participants in the ND group had the lowest number of task failures when using the 0.3 mm key height. Participants in the TB group had the lowest number of task failures when using the 0.7 mm key height, and the highest when using the 0.5 mm key height (Figure 9.3.10). Participants in the MU and SU groups did not commit any task failures when using the different key heights.

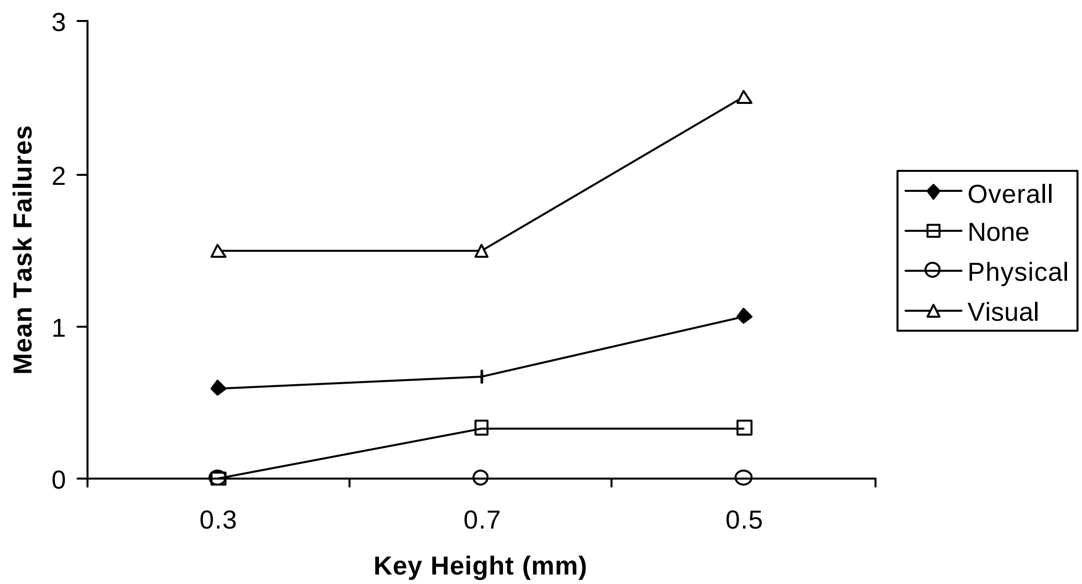


Figure 9.3.10 - Mean Task Failures for All Tasks by Key Height and Disability

Preference Rankings

A two-way analysis revealed significant effects on preference ranking for key height ($F_{(2,24)} = 16.96$, $p < 0.001$), but not for disability, or the key height by disability interaction ($F_{(8,24)} = 1.77$, $p < 0.14$). Subsequent one-way analysis revealed a significant ($p < 0.0001$) difference in overall preference rankings between the 0.5 mm and 0.7 mm key heights, when compared to the 0.3 mm key height (Groups A and B, Figure 9.3.11).

Participants preferred the 0.7 mm key height overall, followed by the 0.5 mm and 0.3 mm key heights. The ND and MU participant groups however, preferred the 0.5 mm key height to the 0.7 mm key height. The 0.3 mm key height was the least preferred for all participant groups.

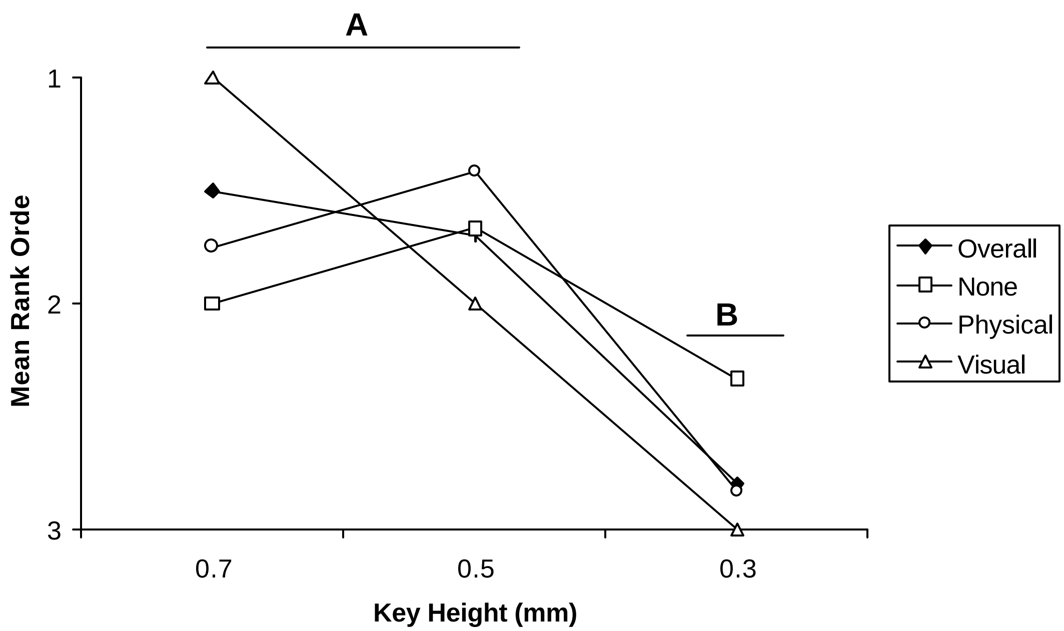


Figure 9.3.11 - Mean Preference Rankings by Key Height and Disability
(Lower numbers indicate greater preference. Groups A and B are significantly different overall.)

Ease of Use Ratings

A two-way analysis of ease of use ratings revealed significant effects for disability ($F_{(4,24)} = 3.77$, $p < 0.04$), and key height ($F_{(2,24)} = 14.77$, $p < 0.0001$), but not for the key height by disability interaction. Subsequent one-way analysis revealed a significant ($p < 0.007$) difference in overall ease of use ratings between the 0.7 mm and 0.5 mm key heights, when compared to the 0.3 mm key height (Groups A and B, Figure 9.3.12).

Participants rated the 0.7 mm key height the highest for ease of use overall, followed by the 0.5 mm and 0.3 mm key heights. This trend was evident for all participant groups, except those in the MU group. Participants in the MU group rated the 0.5 mm key height the highest for ease of use, followed by the 0.7 mm and 0.3 mm key heights.

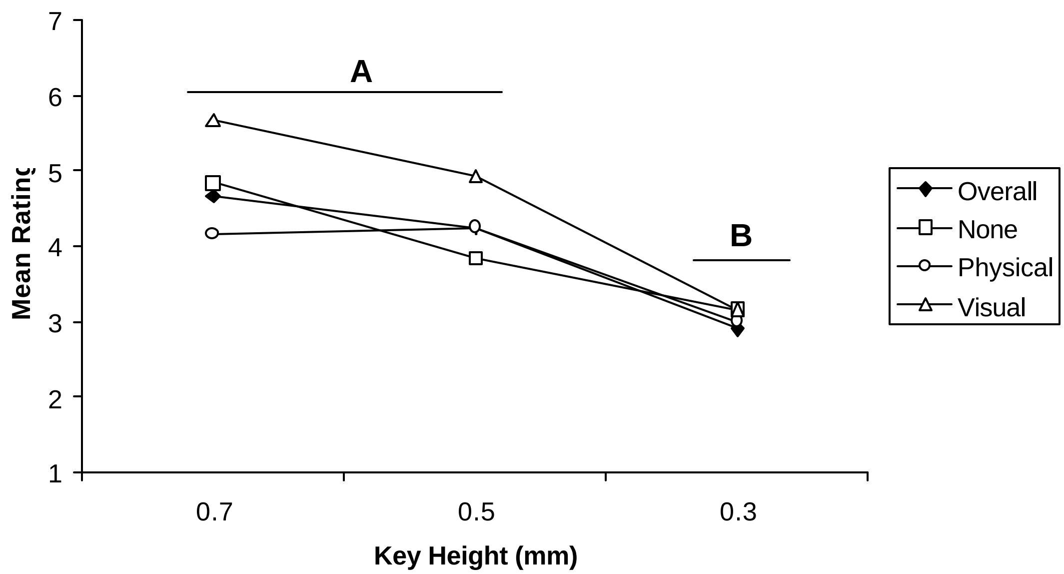


Figure 9.3.12 - Mean Ease of Use Ratings by Key Height and Disability
(Higher numbers indicate greater ease of use rating.)
(Groups A and B are significantly different overall.)

Accessibility Ratings

A two-way analysis of accessibility ratings revealed significant effects for disability ($F_{(4,24)} = 3.12$, $p < 0.07$), and key height ($F_{(2,24)} = 22.94$, $p < 0.0001$), but not for the key height by disability interaction. Subsequent one-way analysis revealed a significant ($p < 0.005$) difference in overall accessibility ratings between the 0.5 mm and 0.7 mm key heights, when compared with the 0.3 mm key height (Groups A and B, Figure 9.3.13).

All participant groups, except for those in the MU group, rated the 0.7 mm key height the highest for accessibility, followed by the 0.5 mm and 0.3 mm key heights. Participants in the MU group rated the 0.5 mm key height the highest for accessibility, followed by the 0.7 mm and 0.3 mm key heights.

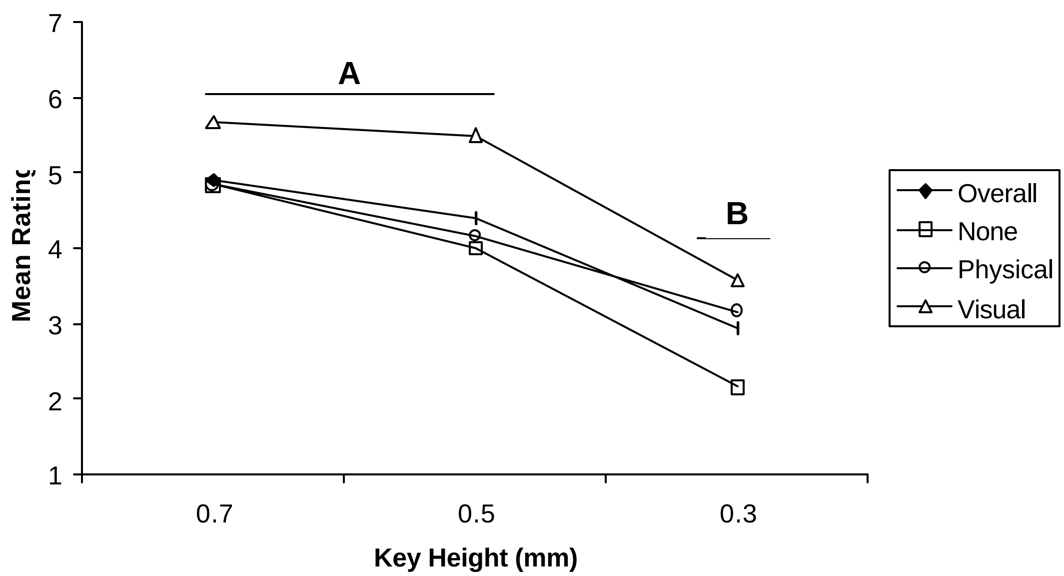


Figure 9.3.13 - Mean Accessibility Ratings by Key Height and Disability
(Higher numbers indicate higher ease of use rating. Groups A and B are significantly different overall.)

9.4 - Effects of <Power> Key Position

Preference Rankings

A two-way analysis of preference rankings for <Power> key position showed significant effects for model ($F_{(7,49)} = 4.86$, $p < 0.0002$), but not for disability or the model by disability interaction. A one-way analysis of overall <Power> key preference rankings also revealed a significant difference in model rankings ($p < 0.0003$).

Participants ranked models 6 and 4 the lowest (most preferred) (Figure 9.4.1). The <Power> keys on these models are relatively large, separate from the numeric keypad, located on the face of the model, but are not molded into the display face. These rankings were significant for participants in the LB ($p < 0.004$), ND ($p < 0.02$), MU ($p < 0.09$), SU ($p < 0.02$), and TB ($p < 0.06$) groups.

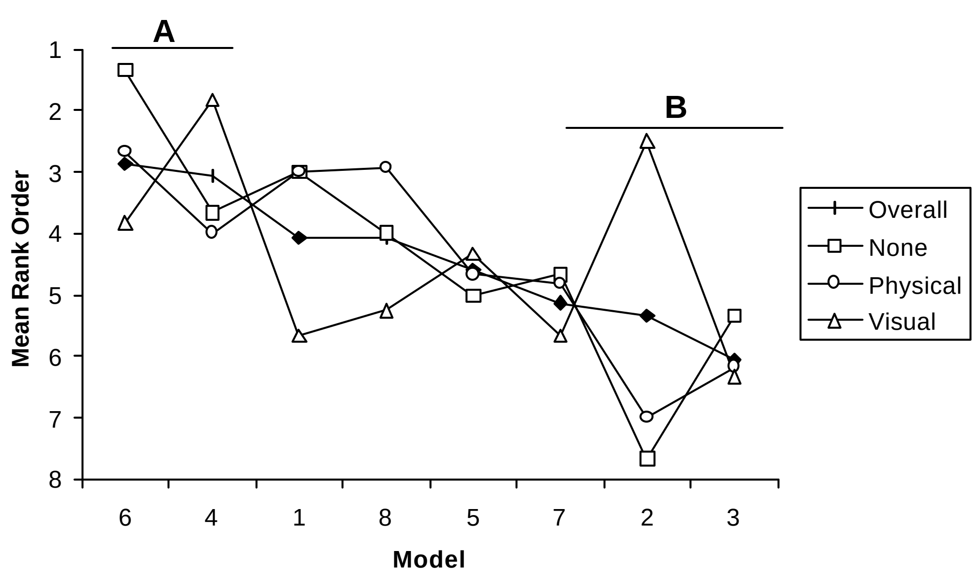


Figure 9.4.1 - Mean Preference Rankings for <Power> Key Positions by Model and Disability
(Lower rank orders indicate greater preference. Groups A and B are significantly different overall.)

Ease of Use Ratings

A two-way analysis of <Power> key ease of use ratings showed significant effects for disability ($F_{(4,49)} = 6.8$, $p < 0.007$) and model ($F_{(7,49)} = 12.88$, $p < 0.0001$), but not for the model by disability interaction. A one-way analysis of overall <Power> key ease of use ratings revealed a significant difference in ratings ($p < 0.0001$), with models 6 and 8 (see Table 7.2.1) rated the highest (Figure 9.4.2).

Participants rated models 6 and 8, which have large <Power> keys located in prominent positions, higher for ease of use than models 2 and 5, which have small <Power> keys that are located in less prominent positions. These ratings were significant for the ND ($p < 0.01$), MU ($p < 0.01$), SU ($p < 0.02$), and TB ($p < 0.06$) groups, but were not significant for the LB group.

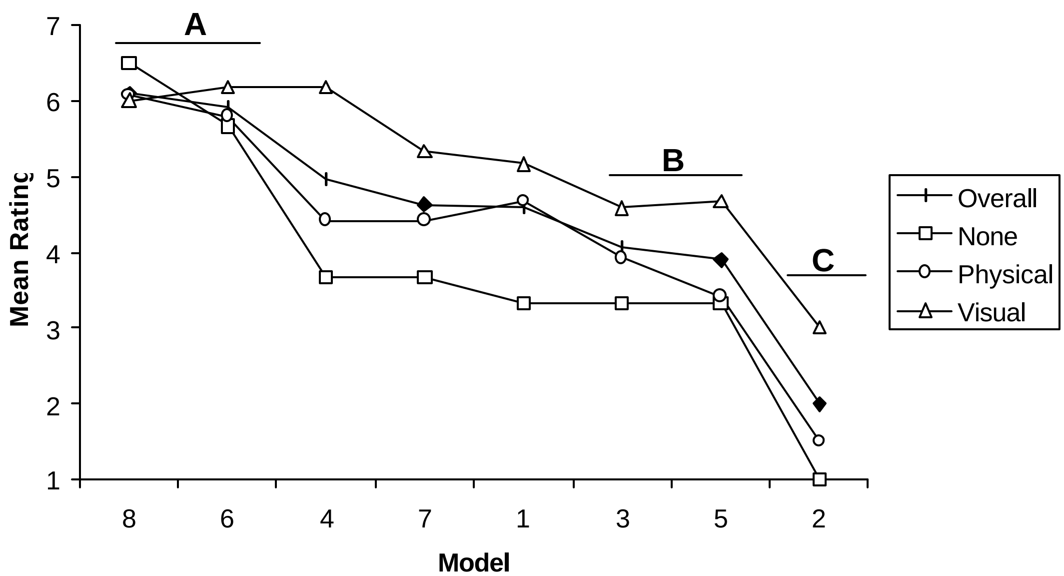


Figure 9.4.2 - Mean Ease of Use Ratings for <Power> Keys by Model and Disability (Higher ratings indicate greater ease of use. Groups A and B are significantly different overall. Group C is significantly different from all models overall.)

Accessibility Ratings

A two-way analysis of accessibility ratings for <Power> key position showed significant effects for disability ($F_{(4,49)} = 6.53$, $p < 0.008$), model ($F_{(7,49)} = 18.54$, $p < 0.0001$), and the model by disability interaction ($F_{(28,49)} = 1.49$, $p < 0.09$). A one-way analysis of overall accessibility ratings was significant ($p < 0.0001$), with models with large <Power> keys located in prominent positions rated higher for accessibility than models with smaller <Power> keys in less prominent locations (Figure 9.4.3).

Accessibility ratings were the highest for models 6 and 8, which have distinct <Power> keys located on the model face, in prominent positions, and away from the numeric keypad. Ratings were significant for the LB ($p < 0.9$), ND ($p < 0.0006$), MU ($p < 0.02$), SU ($p < 0.01$), and TB ($p < 0.008$) groups.

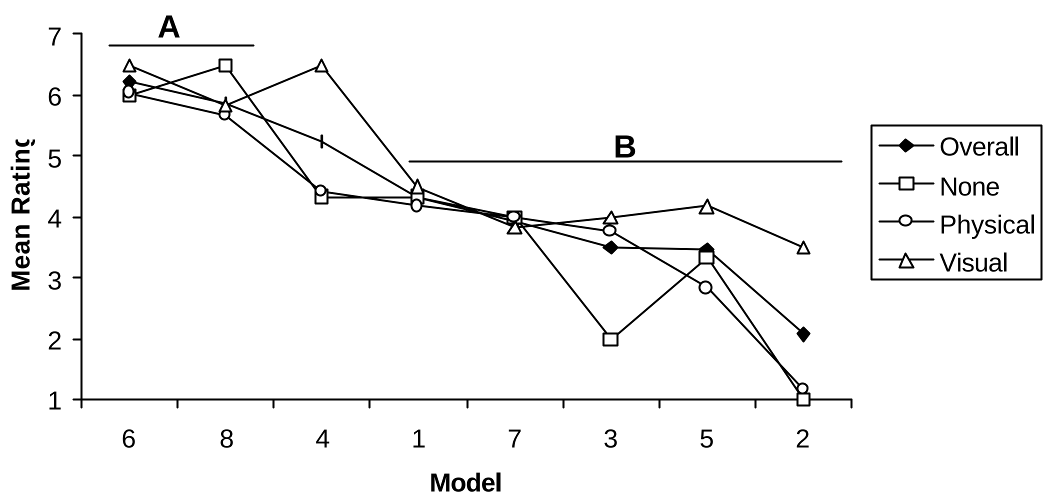


Figure 9.4.3 - Mean Accessibility Ratings for <Power> Keys by Model and Disability
(Higher ratings indicate greater ease of use. Groups A and B are significantly different overall.)

9.5 - Effects of <Send> / <End> Key Position

Preference Rankings

A two-way analysis of preference rankings revealed significant effects for model ($F_{(7,49)}$ = 11.05, $p < 0.0001$), but not for disability or the model by disability interaction. A one-way analysis, overall, also revealed a significant difference in preference rankings ($p < 0.0001$).

Participants preferred large <Send> and <End> keys that were placed in prominent positions, such as above all other keys, or around the multifunction key (i.e. models 6 and 8) versus keys that were small, difficult to distinguish from the numeric keypad, or molded into the display face (models 2, 3, 4, and 5). Preference rankings (Figure 9.5.1) were significant for the LB ($p < 0.0008$), ND ($p < 0.01$), SU ($p < 0.08$), and TB ($p < 0.07$) groups, but not for the MU group.

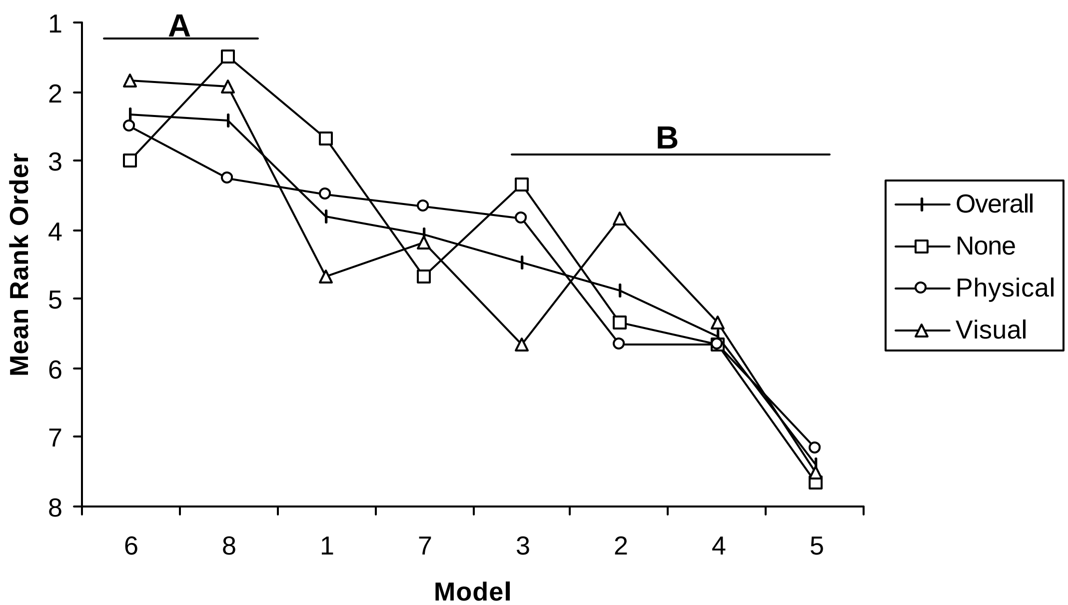


Figure 9.5.1 - Mean Preference Rankings for <Send> / <End> Key Position by Model and Disability (Lower rank orders indicate greater preference. Groups A and B are significantly different overall.)

Ease of Use Ratings

A two-way analysis of ease of use ratings revealed significant effects of model ($F_{(7,49)} = 8.25$, $p < 0.0001$), but not disability or the model by disability interaction. A one-way analysis also revealed a significant difference in ease of use ratings ($p < 0.0001$).

Participants rated models with large <Send> and <End> keys placed in prominent positions significantly higher for ease of use than models with smaller <Send> and <End> keys that were difficult to distinguish from the numeric keypad (Figure 9.5.2). Ease of use ratings were significant for the LB ($p < 0.006$), and TB groups ($p < 0.06$), but not for the ND, MU, and SU groups.

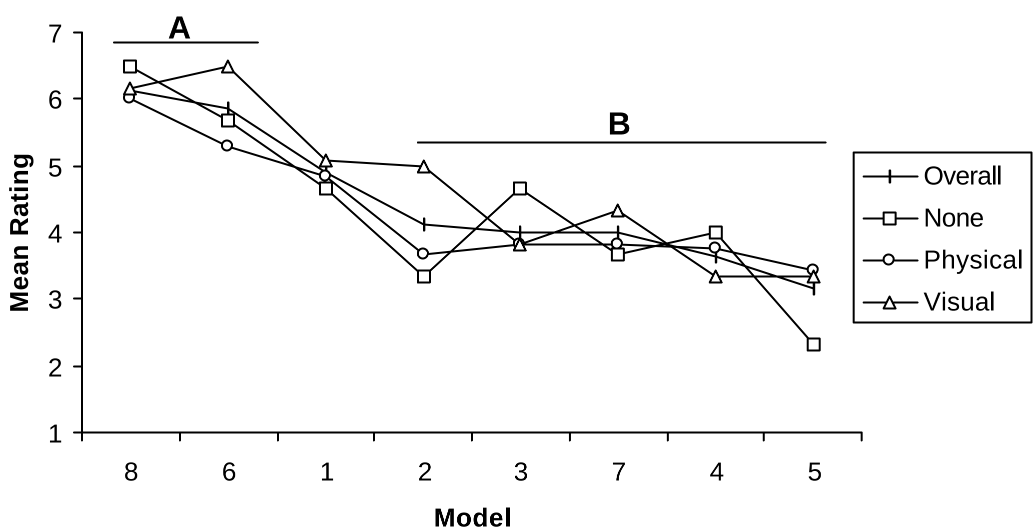


Figure 9.5.2 - Mean Ease of Use Ratings for <Send> / <End> Keys by Model and Disability (Higher ratings indicate greater ease of use. Groups A and B are significantly different overall.)

Accessibility Ratings

A two-way analysis of accessibility ratings revealed significant effects for disability ($F_{(4,49)} = 4.48$, $p < 0.025$) and model ($F_{(7,49)} = 13.46$, $p < 0.0001$), but not for the model by disability interaction ($F_{(28,49)} = 1.36$, $p < 0.15$). A subsequent one-way analysis, revealed a significant difference in the <Send> and <End> key accessibility ratings overall ($p < 0.0001$).

Participants rated large <Send> and <End> keys, placed in prominent positions as the most accessible. The accessibility ratings for models with these features, such as models 6 and 8, were higher than the accessibility ratings for models with smaller <Send> and <End> keys, placed in less prominent positions (Figure 9.5.3). Ratings were significant for participants in the ND ($p < 0.01$), MU ($p < 0.02$), and TB ($p < 0.0005$) groups, but not for participants in the LB ($p < 0.11$) and SU groups.

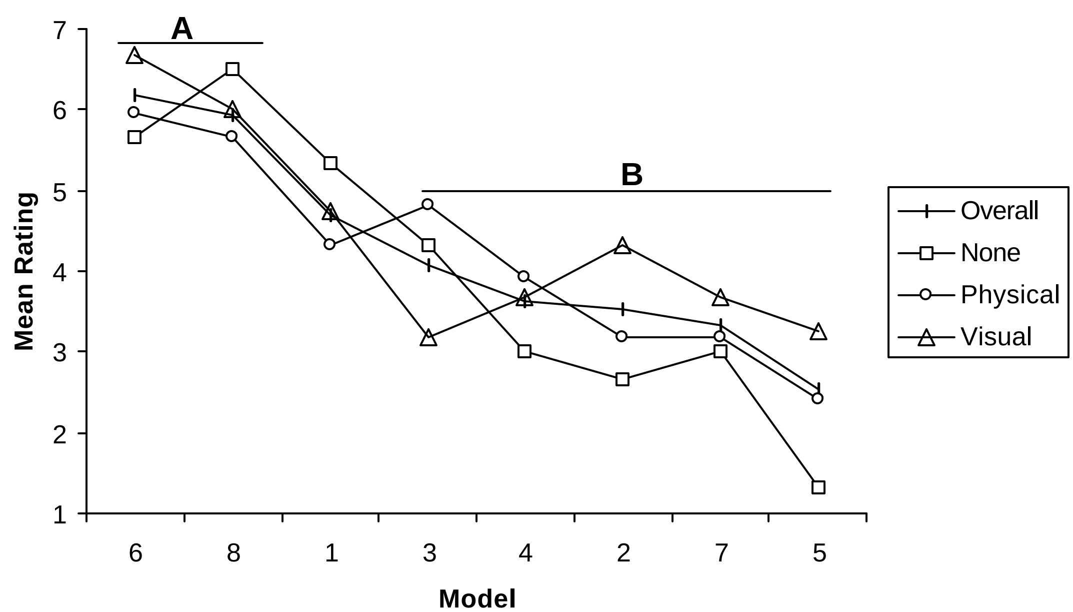


Figure 9.5.3 - Mean Accessibility Ratings for <Send> / <End> Keys by Model and Disability (Higher ratings indicate greater accessibility. Groups A and B are significantly different overall.)

9.6 - Effects of Numeric Font Size

Preference Rankings

A two-way analysis of numeric font size preference rankings revealed significant effects for numeric font size ($F_{(5,31)} = 6.58$, $p < 0.0001$), but not for disability, or the numeric font size by disability interaction. One-way analysis of overall preference rankings indicates a significant difference in preference ratings, with participants preferring the 22-point and 36-point fonts significantly more than the 16-point font ($p < 0.0001$). There was not a significant difference in preference for normal or bold font type, however (Figure 9.6.1).

Participants in the LB group showed a significant difference in preference between the 22-point and 36-point fonts, when compared to the 16-point fonts ($p < 0.002$). The difference in preference between normal and bold fonts for the LB group was not significant. Participants in the ND and MU groups did not strongly favor one font size over another, and there were no significant differences in the rankings for these participants. Those in the SU participant group preferred the 36-point font, followed by the 22-point and 16-point fonts. These participants generally preferred the bold fonts to the normal fonts. The preference for the 36-point fonts, as well as the 22-point bold font among those in the SU participant group was significantly different than the preference for the normal weight 22-point font, and both the normal and bold 16-point fonts ($p < 0.0001$).

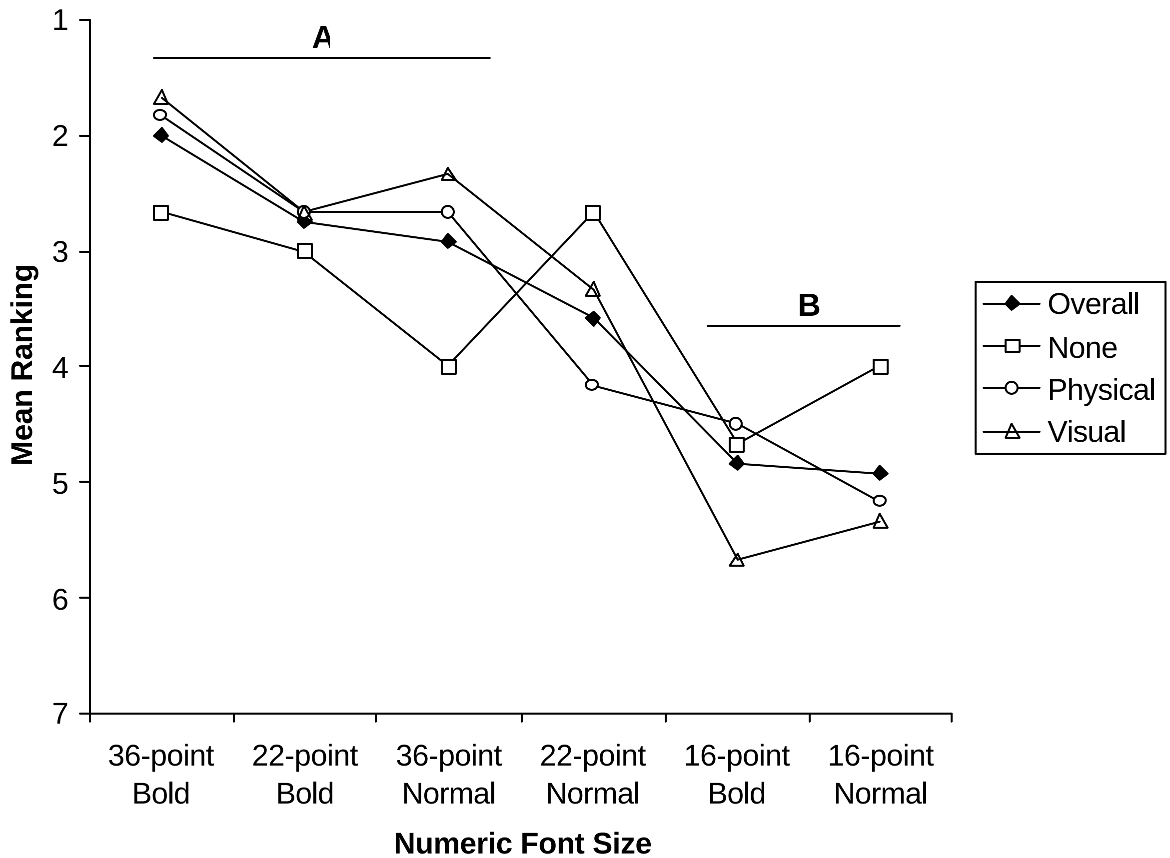


Figure 9.6.1 - Mean Preference Rankings by Numeric Font Size and Disability
(Lower numbers indicate greater preference. Groups A and B are significantly different overall.)

Ease of Use Ratings

A two-way analysis of overall ease of use ratings revealed significant effects for numeric font size ($F_{(5,31)} = 10.44$, $p < 0.0001$), but not for disability ($F_{(3,31)} = 2.84$, $p < 0.11$), or the numeric font size by disability interaction. One-way analyses of overall ease of use ratings revealed that participants rated the 22-point and 36-point fonts significantly higher for ease of use than the 16-point fonts ($p < 0.0001$). This trend, although not significant, was evident for participants in the ND, and MU groups. Participants in the LB group rated the 36-point fonts significantly higher for ease of use than the 16-point fonts ($p < 0.0006$), while participants in the SU group rated the 36-point fonts significantly higher than the 22-point fonts, and the 22-point fonts significantly higher than the 16-point fonts ($p < 0.0001$). There was no significant difference in ease of use ratings between the normal fonts, and the bold fonts (Figure 9.6.2).

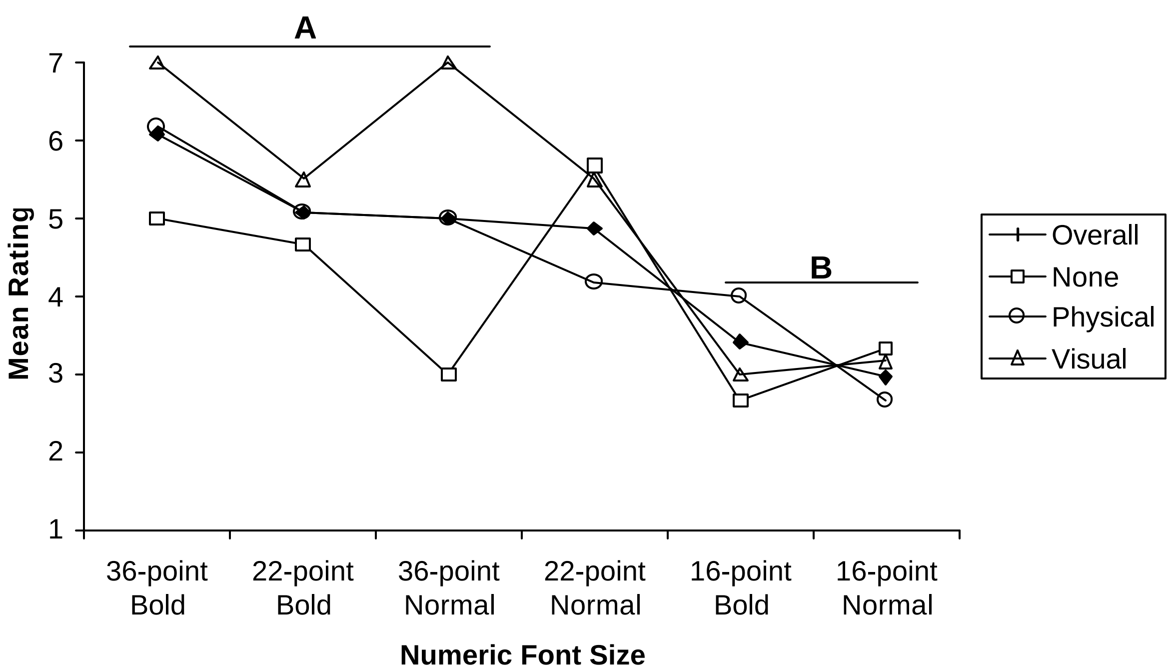


Figure 9.6.2 - Mean Ease of Use Ratings by Numeric Font Size and Disability
(Higher ratings indicate greater ease of use. Groups A and B are significantly different overall.)

Accessibility Ratings

A two-way analysis of accessibility ratings showed significant effects of numeric font size ($F_{(5,31)} = 6.69$, $p < 0.0001$), disability ($F_{(3,31)} = 2.92$, $p < 0.10$), and the numeric font size by disability interaction ($F_{(15,31)} = 1.86$, $p < 0.06$). One-way analyses, overall, revealed that participants rated the 22-point and 36-point fonts significantly higher for accessibility than the 16-point fonts ($p < 0.002$). Participants in the LB group rated the 22-point bold font the highest for accessibility, while those in the ND participant group rated the 22-point normal font the highest for accessibility. Both groups rated the 16-point bold font the lowest for accessibility. Participants in the MU group rated the 36-point fonts the highest for accessibility, followed by the 22-point fonts. Differences in font size ratings for the LB, ND, and MU groups were not significant. Participants in the SU group rated the 36-point fonts significantly higher for accessibility than the 22-point fonts, and these fonts significantly higher than the 16-point fonts ($p < 0.0001$) (Figure 9.6.3).

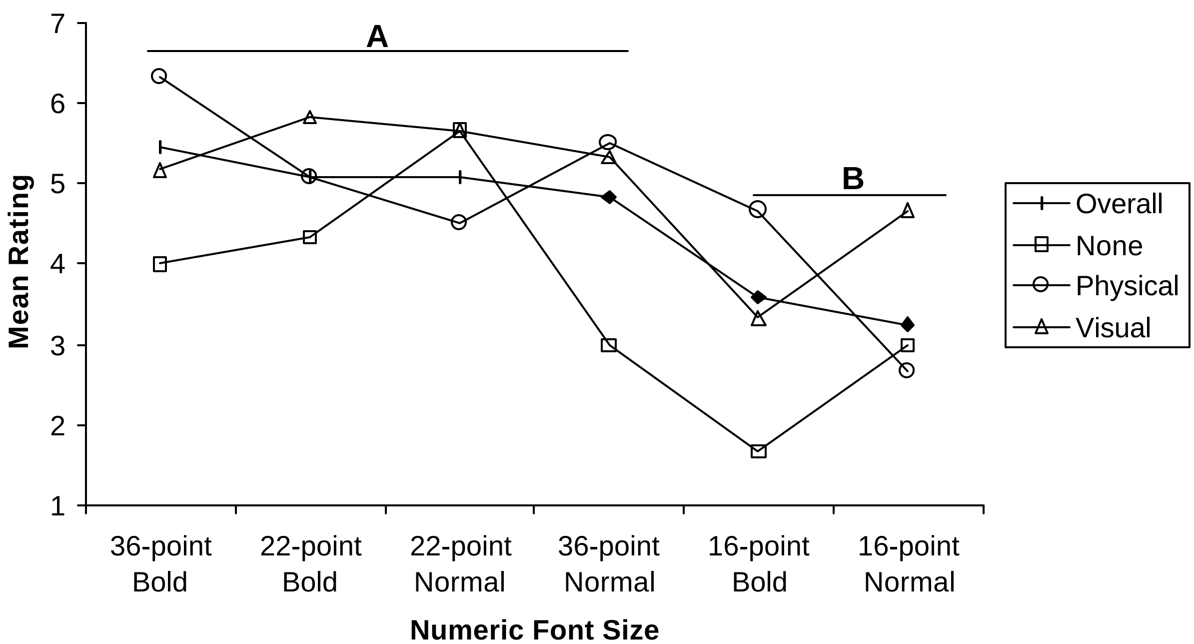


Figure 9.6.3 - Mean Accessibility Ratings by Numeric Font Size and Disability
(Higher ratings indicate greater accessibility. Groups A and B are significantly different overall.)

9.7 - Effects of Alphabetic Font Size

Preference Rankings

A two-way analysis of overall preference rankings revealed significant effects for alphabetic font size ($F_{(2,19)} = 15.8$, $p < 0.0002$), but not for disability or the font size by disability interaction. One-way analyses, overall, revealed that participants had a greater preference for the 22-point font than the 16-point and 36-point fonts, with the 22-point and 36-point fonts being preferred significantly more than the 16-point fonts ($p < 0.0001$). This trend was evident across all participant groups (Figure 9.7.1).

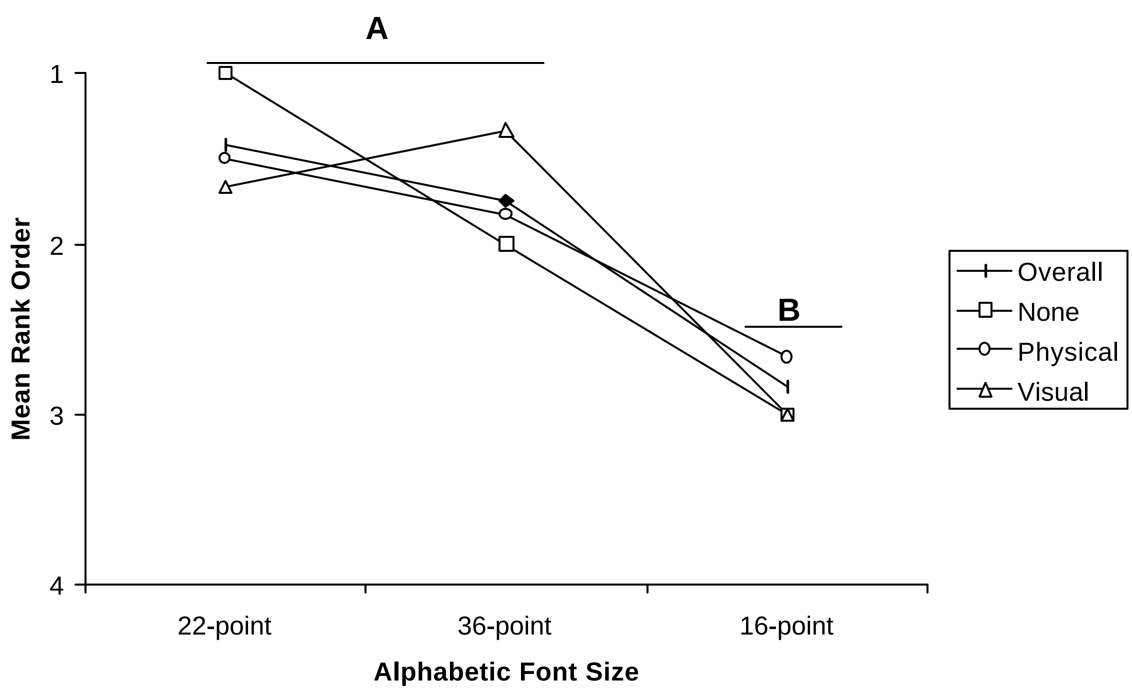


Figure 9.7.1 - Mean Preference Rankings by Alphabetic Font Size and Disability
(Lower rankings indicate greater preference. Groups A and B are significantly different overall.)

Ease of Use Ratings

A two-way analysis of ease of use ratings revealed significant effects for alphabetic font size ($F_{(2,19)} = 42.04$, $p < 0.0001$), and the font size by disability interaction ($F_{(6,19)} = 3.87$, $p < 0.01$), but not for disability. Subsequent one-way analysis revealed that participants rated the 22-point and 36-point fonts the highest for ease of use overall, with ratings significantly higher than those for the 16-point font ($p < 0.0001$). Participants in the LB and SU groups rated the 36-point alphabetic fonts significantly higher for ease of use than the 16-point and 22-point size alphabetic fonts ($p < 0.001$ and $p < 0.05$, respectively). Participants in the ND group rated the 22-point font the highest for ease of use, followed closely by the 36-point font. These ratings were significantly higher than those for the 16-point font ($p < 0.02$). Differences in the font size preference for participants in the MU group were not significant (Figure 9.7.2).

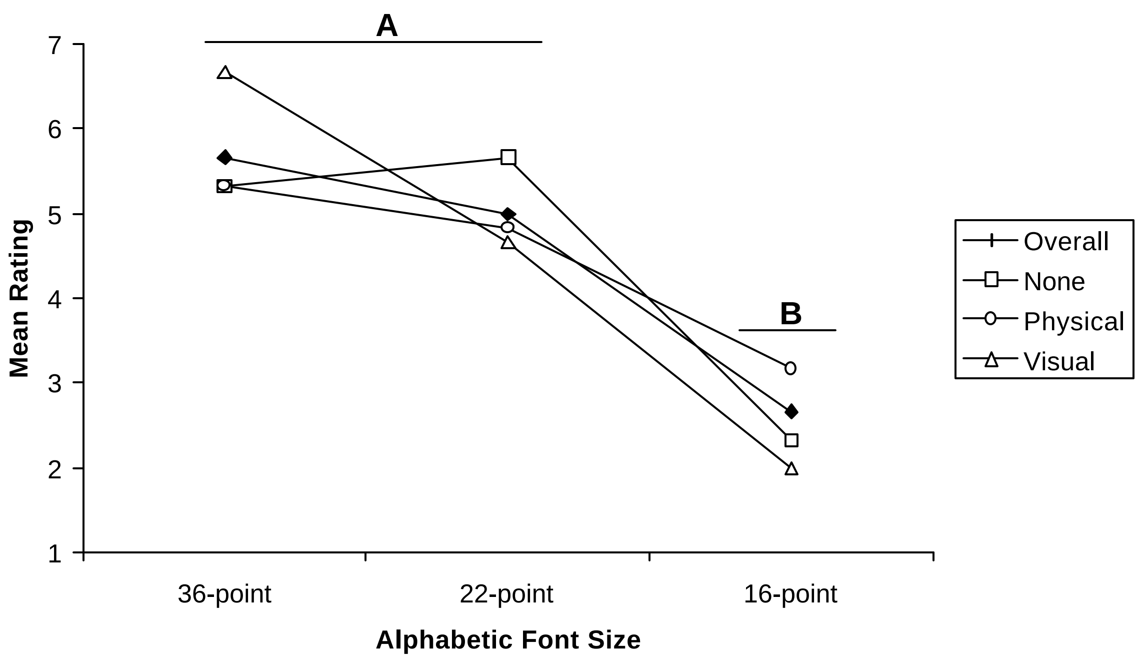


Figure 9.7.2 - Mean Ease of Use Ratings by Alphabetic Font Size and Disability
(Higher ratings indicate higher ease of use. Groups A and B are significantly different overall.)

Accessibility Ratings

A two-way analysis of accessibility ratings revealed significant effects for alphabetic font size ($F_{(2,19)} = 39.34$, $p < 0.0001$), and the alphabetic font size by disability interaction ($F_{(6,19)} = 3.29$, $p < 0.03$), but not for disability. A one-way analysis revealed that participants rated the 22-point and 36-point fonts significantly higher for accessibility than the 16-point font ($p < 0.0001$). Participants in the LB and ND groups rated the 22-point font significantly higher than the 16-point font for accessibility ($p < 0.005$ and $p < 0.01$, respectively). Participants in the SU group rated the accessibility of the 22-point and 36-point fonts significantly higher than the 16-point font ($p < 0.008$). Participants in the MU group rated the three font sizes similarly for accessibility (Figure 9.7.3).

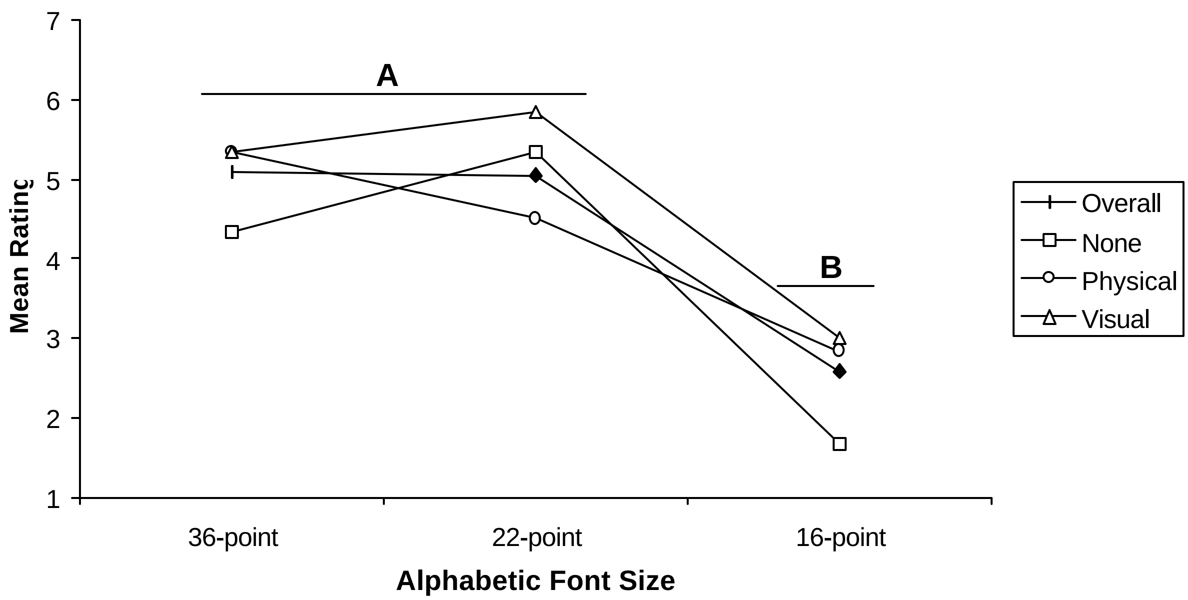


Figure 9.7.3 - Mean Accessibility Ratings by Alphabetic Font Size and Disability
(Higher ratings indicate greater accessibility. Groups A and B are significantly different overall.)

9.8 - Results Summary

To facilitate later recommendations of specific levels of each keypad and display attribute, each level of independent variable was rank ordered using the overall results from each of the different dependent measures (7-digit time, 7-digit error, preference ranking, etc.). The ranks for each level across the dependent measures were then added. The lowest sum of ranks can thereby be considered the “best” across the different dialing and subjective measures. Rank sums for lateral pitch (Table 9.8.1), keystroke (Table 9.8.2), key height (Table 9.8.3), <Power> key (Table 9.8.4), <Send> / <End> key (Table 9.8.5), numeric font size (Table 9.8.6), and alphabetic font size (Table 9.8.7) are shown below.

Table 9.8.1 - Overall Rank Order Summary by Lateral Pitch

Lateral Pitch	Time						Task Failures						Subjective Measures				Total Rank Sum
	7-Digit	11-Digit	SMR	911	Total Time	Rank Sum	7-Digit	11-Digit	SMR	911	Total TF	Rank Sum	Pref.	Ease of Use	Acc.	Rank Sum	
10 mm	4	4	3	2	4	17	4	4	4	1	4	17	4	4	4	12	46
11 mm	3	3	2	3	3	14	3	3	2.5	2.5	3	14	3	3	3	9	37
12 mm	1	1	1	1	1	**5**	1	1	1	4	1	**8**	1	1	1	**3**	**16**
13 mm	2	2	4	4	2	14	2	2	2.5	2.5	2	11	2	2	2	6	31

Table 9.8.2 - Overall Rank Order Summary by Keystroke

Keystroke	Time						Task Failures						Subjective Measures				Total Rank Sum
	7-Digit	11-Digit	SMR	911	Total Time	Rank Sum	7-Digit	11-Digit	SMR	911	Total TF	Rank Sum	Pref.	Ease of Use	Acc.	Rank Sum	
0.3 mm	3	2	1	3	2	11	2.5	2	2	1	2	9.5	3	2	1.5	6.5	27
0.5 mm	2	3	2	2	3	12	1	1	2	2	1	**7**	1	3	3	7	26
0.7 mm	1	1	3	1	1	**7**	2.5	3	2	3	3	13.5	2	1	1.5	**4.5**	**25**

Table 9.8.3 - Overall Rank Order Summary by Key Height

Key Height	Time						Task Failures						Subjective Measures				Total Rank Sum
	7-Digit	11-Digit	SMR	911	Total Time	Rank Sum	7-Digit	11-Digit	SMR	911	Total TF	Rank Sum	Pref.	Ease of Use	Acc.	Rank Sum	
0.3 mm	3	3	3	3	3	15	2	1	1.5	1.5	1	**7**	3	3	3	9	31
0.5 mm	2	2	2	2	2	10	3	2	3	3	3	14	2	2	2	6	30
0.7 mm	1	1	1	1	1	**5**	1	3	1.5	1.5	2	9	1	1	1	**3**	**17**

Table 9.8.4 - Overall Rank Order Summary by <Power> Key

#	Model	Preference	Ease of Use	Accessibility	Rank Sum
1	Audiovox CDM-9100	3	5	4	12
2	Nokia 8260	7	8	8	23
3	Ericsson T28	8	6	6	20
4	Motorola V2260	2	3	3	8
5	Motorola Star-Tac	5	7	7	19
6	**Audiovox CDM-4000**	**1**	2	**1**	**4**
7	Sanyo SCP-5000	6	4	5	15
8	Audiovox CDM-9000	4	**1**	2	7

Table 9.8.5 - Overall Rank Order Summary by <Send> / <End> Key

#	Model	Preference	Ease of Use	Accessibility	Rank Sum
1	Audiovox CDM-9100	3	3	3	9
2	Nokia 8260	6	4	6	16
3	Ericsson T28	5	6	4	15
4	Motorola V2260	7	7	5	19
5	Motorola Star-Tac	8	8	8	24
6	**Audiovox CDM-4000**	**1**	2	**1**	**4**
7	Sanyo SCP-5000	4	5	7	16
8	Audiovox CDM-9000	2	**1**	2	5

Table 9.8.6 - Overall Rank Order Summary by Numeric Font Size

Numeric Font	Preference	Ease of Use	Accessibility	Rank Sum
16 Pt. Normal	6	6	6	18
16 Pt. Bold	5	5	5	15
22 Pt. Normal	4	4	2.5	10.5
22 Pt. Bold	2	2	2.5	6.5
36 Pt. Normal	3	3	4	10
36 Pt. Bold	**1**	**1**	**1**	**3**

Table 9.8.7 - Overall Rank Order Summary by Alphabetic Font Size

Alphabetic Font	Preference	Ease of Use	Accessibility	Rank Sum
16 Pt.	3	3	3	9
22 Pt.	**1**	2	2	5
36 Pt.	2	**1**	**1**	**4**

The lowest sum of ranks for each level of independent variable can provide designers with an indication as to what level of each independent variable should be produced in a cellular telephone model. In some instances, the rank sums clearly demonstrate that one level of independent variable should be more highly considered than the others. In other cases however, the rank sums may be too close to reliably consider one level of independent variable over the others. Designers should use these rank sums with some caution, however, and incorporate other information not considered in this study, such as knowledge of the market, production costs, and availability when making a determination on which level to choose.

The 12 mm lateral pitch was clearly ranked the best overall, since the rank order was the lowest in all but one of the measures. The 13 mm lateral pitch was ranked a distant second, followed by the 11 mm and 10 mm levels, respectively.

The 0.7 mm level of keystroke received the lowest rank sum. However, due to the minimal differences in rank sums between each keystroke level, the choice of which level to design should not be based solely on study results. Study results should be used in conjunction with other information, such as production cost, availability, etc. The decision of which level to use in instances such as this, where the difference in rank sums is minimal, may be based on specific criteria such as error tolerance, since excessive errors may reduce user satisfaction. If error tolerance is used as the decision criteria, the 0.5 mm keystroke should be chosen, since the rank sum for the number of task failure-related measures was the lowest.

The 0.7 mm level of key height was ranked the best overall, since the rank order was the lowest in all but the task failure-related measures. The 0.5 mm key height was ranked a distant second, followed by the 0.3 mm key height, which ranked last in all but the task failure-related measures.

The rank orders of the <Power>, <Send> and <End> key measures indicate that models with large, distinct <Power>, <Send> and <End> keys that are located on the model face, separate from the numeric keypad, and not part of the display face scored the best overall. This was demonstrated by the low rank sums for the Audiovox model CDM-4000 and CDM-9000. Both of these models have large, distinct <Power>, <Send> and <End> keys that were separate from the numeric keypad, but not molded into the display face. Although models CDM-4000 and CDM-9000 had the best overall rank scores, these scores were not endorsements of those particular models. Instead, participants were asked to rank and rate these keys solely on position, and how easy they were to use, not on the rest of the model. These models merely demonstrated possible key positions and sizes for participants to evaluate.

The rank sum for the numeric fonts was the lowest for the 36-point bold font, followed closely by the 22-point bold font. The rank sums for the 22-point and 36-point normal fonts, although higher in rank sums than the bold fonts, were very close. The rank sums for the alphabetic font, while only evaluated with the normal weight, followed the same trend, with the 36-point font ranked the best, followed very closely by the 22-point font.

Careful judgment is required when deciding which size and weight of font to select. Although eight out of twelve participants indicated a preference for a larger font size versus the amount of information displayed, some only preferred a larger font size if the telephone did not have special capabilities, such as text messaging or electronic mail. These participants wanted a compromise between font size and the amount of information displayed for telephones with these features. Designers must be cognizant of the features of each telephone when making this decision.

10.0 - Discussion – Cellular Telephone Study

The intent of this study was to determine what keypad and display font attributes provided the fastest task completion time, lowest number of task failures, and highest subjective evaluation measures among participants with visual and upper extremity physical disabilities. This study was a continuation of a previous survey-based Smith-Jackson et. al (2001) study, where participants with disabilities indicated concerns over cellular telephone keypad and display designs.

Participants completed common dialing tasks, wherein the time to complete the task, and the number of task failures was quantified. After the dialing tasks were completed, participants rank ordered telephone models based on the attribute of interest, and then rated each attribute on ease of use and accessibility.

10.1 - Hypotheses

It was hypothesized that main effects and interaction effects of key attributes on task completion time, task failures, and subjective evaluation measures among the different disability groups would be present. Additionally, it was hypothesized that there would be main effects and interaction effects of the display font size, as well as the <Power>, <Send>, and <End> key locations on subjective evaluations.

A main effect of disability was present with all of the keypad attributes evaluated during the dialing tasks. In several instances, task completion times for all disability groups followed a similar trend, however, participants with disabilities generally required longer to complete the tasks. Participants with disabilities did not commit nearly as many task failures as expected; they just completed tasks more slowly.

Main effects and interaction effects of lateral pitch and disability were present in nearly all of the dialing tasks. Participants in all disability groups generally completed dialing tasks more quickly, and with less task failures when using the 12 mm lateral pitch. Task completion times and the number of task failures with the 13 mm lateral pitch were lower than those with the 10 mm and 11 mm models. These main effects and interaction effects indicate that performance (task completion times and error rate) was the best when using the larger models, regardless of disability.

Main effects and interaction effects for lateral pitch and disability for the preference, ease of use, and accessibility measures followed trends similar to these for the dialing tasks. Participants ranked and rated telephone models in a similar manner. Participants with visual disabilities generally rated models better than the other participant groups.

The higher ratings for those with visual disabilities may be due to a "Pollyanna phenomenon" which is used as a strategy to cope with disabilities. Saeterdsal (1997) believes those with disabilities may believe that "every cloud has a silver lining" and tend to rate events in life more highly than those without disabilities. It is unclear why those with visual disabilities, and not those with upper extremity disabilities exhibited this effect, however.

Main effects and interaction effects due to keystroke and disability type were not apparent in many of the dialing tasks. Although participants completed tasks the most quickly with the 0.7 mm keystroke overall, and the fewest task failures with the 0.5 mm keystroke overall, it is unclear whether performance was due to keystroke, or if it was a random occurrence.

There were no main effects and interaction effects for keystroke and disability for the preference, ease of use, and accessibility measures. It is important to note that although there were no main effects or interaction effects, participants with visual disabilities continued to rate each level of independent variable better for ease of use and accessibility than the other participant groups.

Main effects and interaction effects due to key height and disability type were present in only a portion of the dialing tasks. The main effect of key height was apparent at the 0.7 mm and 0.3 mm levels. Task completion times were the lowest with the 0.7 mm key height, while the number of task failures was the lowest with the 0.3 mm key height. The main effect of disability was relatively consistent across all the different task types, with those with no disabilities completing tasks significantly faster than participants in the other participant groups. Task completion times for the visual and physical disability groups were very similar. Interaction effects did not appear frequently for the task completion times, however they did appear frequently in the task failures.

The number of task failures was the highest when participants with disabilities used the 0.5 mm key height. The reasons for this, however, were unclear.

Main effects and interaction effects for key height and disability were present for the preference, ease of use, and accessibility measures. The 0.7 mm key height scored the best for the overall preference rankings, and for participants with visual disabilities. Participants with physical disabilities and no disabilities, however, scored the 0.5 mm key height greater for preference. The consistent preference rankings of participants with visual disabilities outweighed the preference rankings of the other participants, leading to the highest overall rating for the 0.7 mm key height. Ease of use and accessibility ratings followed similar trends, with participants generally rating the 0.7 mm key height the highest. Participants with visual disabilities continued to rank the levels of key height consistently better than other participants.

A main effect of numeric display font size was present for the preference, ease of use, and accessibility ratings. Nearly all users rated the 22-point and 36-point fonts more highly for each of the measures than the 16-point font. A main effect of disability was not present for the preference or ease of use ratings. Nearly all participants, regardless of disability, rated the larger fonts more highly than the 16-point font. Interaction effects of display font and disability were present for the accessibility ratings. Participants with physical disabilities rated the 16-point bold font much higher for accessibility than those with no disabilities (mean ratings were 4.67 and 1.67, respectively).

Main effects for alphabetic font size were present for display font size for the preference, ease of use, and accessibility ratings. Nearly all participants ranked the 22-point and 36-point fonts higher than the 16-point font in each of these measures. Main effects due to the alphabetic font size by disability interaction were present in the ease of use and accessibility ratings. Although an interaction was present, graphs for ease of use and accessibility (Figures 9.7.2 and 9.7.3) reveal little difference in the trends between disabilities for each level of display font. This interaction, while present quantitatively, may not have any qualitative or practical implications.

Main effects for <Power> key position were present for the preference, ease of use, and accessibility ratings. Models with large, distinct <Power> keys that were separate from the numeric keypad were rated more highly by all participant groups than models with small <Power> keys, or <Power> keys located within the numeric keypad.

Main effects for disability were present for the <Power> key ease of use ratings. The ease of use ratings for each model (<Power> key) followed similar trends, although the rating levels differed between each disability group. Participants with visual disabilities rated all models, except models 6 and 8, more highly for ease of use than other participants. Participants with physical disabilities rated most models, except models 6 and 8, similarly to those with visual disabilities, though the ratings were lower. Those with no disabilities rated all models, except models 5, 6, and 8, similarly to the other groups, though again the ratings were lower. Ratings for models 6 and 8 were relatively similar between participant groups.

Main effects and interaction effects for the <Power> key accessibility ratings were present for model (<Power> key), disability, and the model by disability interaction. Model ratings were relatively similar for each participant group, except for models 3 and 4. Participants with visual disabilities rated model 4 more highly than other participant groups. Those with no disabilities rated model 3 much lower than other participant groups. These differences in ratings between the two disability groups for the different models resulted in significant quantitative effects, however qualitative effects and practical relevance may not be important.

The main effect of model (<Send> and <End> key position) was present in the <Send> and <End> key preference ranking and ease of use ratings. Participants rated models 6 and 8 the highest for preference and ease of use, regardless of disability group. Slight variations in rankings and ratings were present for each disability group, however these variations were not significant.

Main effects for model (<Send> and <End> key position), and disability were present in the <Send> and <End> key accessibility ratings. Ratings for models 6 and 8 were significantly higher than ratings for models 2, 3, 4, 5, 7. Participants with no disabilities rated models 1 and 8 more highly for accessibility, and rated models 2, 4, 5, and 7 less than other participant groups for accessibility. Participants with visual disabilities rated models 2, 5, and 7 more highly for accessibility, and rated model 3 lower than other groups for accessibility.

10.2 - Interpretation of Results

Lateral Pitch

The fastest overall task completion times, and lowest overall number of task failures occurred using the 12 mm lateral pitch. However, depending on the task and disability category, the task completion times and the number of task failures for the 13 mm lateral pitch were the lowest. Performance with the 13 mm lateral pitch, though, was not consistent for specific disability categories across all task types. In many cases, differences in task completion time and the number of task failures between the 12 mm and 13 mm lateral pitches were not significant, and usually very small.

Task completion times and the number of task failures were the highest when using the 10 mm and 11 mm models. Although the distance the finger must travel between keys on these models was small, more precision was needed to ensure the proper target (key) was being selected (Sanders and McCormick, 1993; Wickens, 2000). The 11 mm telephone was also used to test the different levels of key height, and these different key heights may have slowed the performance of participants that dialed by touch rather than sight.

Subjective evaluation scores were high for the 12 mm and 13 mm models, and low for the 10 mm and 11 mm models. The differences in scores may have been confounded due to other factors, such as the size of each model, the separation of the keys, and the exterior finish. Participants were asked to exclude considerations other than lateral pitch in these evaluations, but these considerations may still have affected the results.

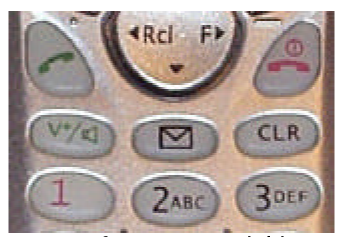
Although factors other than lateral pitch may have influenced the preferences for the 12 mm and 13 mm lateral pitches, these results are consistent with a study of standard and reduced sized numeric keypads by Loricchio (1991). In this study, participants preferred keypads with a greater lateral pitch (center-to-center key spacing) over keypads with a smaller lateral pitch.

Both the 12 mm and 13 mm models seemed large enough to fit participants' hands relatively well, however they were not so large that participants became uncomfortable with them. The 10 mm model was light and compact, and difficult to hold for several users, while the 11 mm model was slightly larger, and not as difficult to hold.

The keys on the 12 mm and 13 mm models were slightly larger and more distinct than those on the 10 mm and 11 mm models. The 10 mm and 11 mm models used keys that had little separation, and were relatively similar in size and shape for both the numeric keypad, and special function keys such as <Messages>, <Clear>, <Send>, and <End>. Several participants committed task failures when trying to press the <1>, <2>, and <3> keys on both the 10 mm and 11 mm models, often mistaking these keys for the similarly sized and shaped keys in the row above. Additionally, the keys on the 10 mm models, and one of the 11 mm models were flush with the case. Participants with visual disabilities may have had difficulty differentiating some of these keys, due to similar size, shape, protrusion, and proximity to the numeric keypad (Figure 10.2.1). Difficulty differentiating numeric keys and special function keys did not seem apparent when using the 12 mm and 13 mm models.

According to Sanders and McCormick (1993), the location, spacing, and shape of keys can provide important information regarding the function of the key. The 12 mm and 13 mm models used location, spacing, and shape coding to differentiate the <Power>, <Send>, and <End> keys from the numeric keypad, whereas the 10 mm and 11 mm models did not.

Top rows of 10 mm model keypad.



Top rows of 11 mm model keypad.

Figure 10.2.1 - Key Size and Shape Similarities on the 10 mm and 11 mm Models

The exterior case finish of the 12 mm and 13 mm models was such that the telephones did not slip from the participants grasp. The exterior finish on the 10 mm model was very smooth and difficult to grip. Many participants committed task failures by dropping the 10 mm model during the study, and the combination of small size and finish may have contributed to this. The finish on the 11 mm model was not as smooth, and participants had an easier time gripping it.

According to Bailey (1993), the decision to select design features should be based on a composite of the objective and subjective measures for specific features. Objective measures include task completion time and number of task failures, while subjective measures include preference, ease of use, and accessibility ratings. The 12 mm lateral pitch was the best in each of these measures, followed closely by the 13 mm lateral pitch. It is interesting to note that the task completion times and the number of task failures increased, and subjective measures decreased, for the 10 mm and 11 mm models as the lateral pitch decreased (Figure 10.2.2).

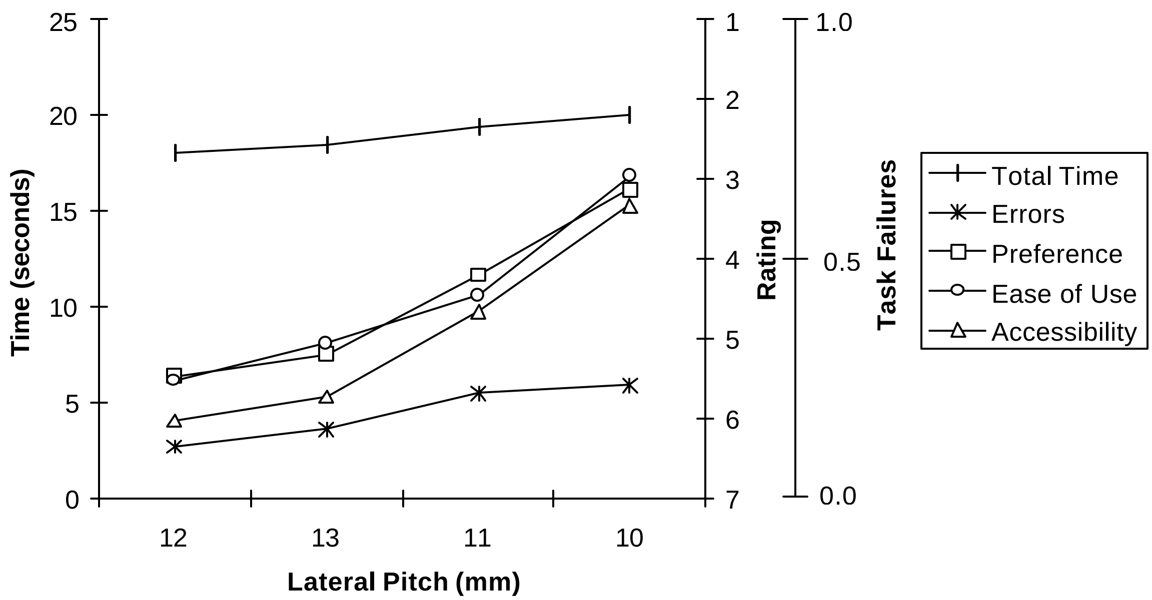


Figure 10.2.2 – Mean Overall Objective and Subjective Measures by Lateral Pitch
(Y-axes are oriented so that values closer to the origin indicate better results.)

Keystroke

Differences in task completion times due to keystroke were small, and not significant. The number of task failures differed between levels, and were generally the lowest when using the 0.5 mm keystroke. In many cases the difference in the number of task failures was significant.

Participants in the SU group had a higher number of task failures when using the 0.7 mm keystroke. It is unclear whether the number of task failures can be attributed to key design, or fatigue. Participants committed several task failures by brushing the top of a key, which may have been inadvertently activated due to lateral key movement (Figure 10.2.3). Activation in this manner seemed more frequent with the 0.7 mm level of keystroke compared to the other levels. Participants in the SU group received the 0.7 mm level of lateral pitch in the latter half of the presentation order (presentations 7, 8, and 9), when they may have been fatigued. A partial counterbalance of presentation order should have prevented this if participants in each group were evenly distributed throughout the study. Since participants in the SU group were difficult to recruit, the order in which they participated occurred when the presentation order of the 0.7 mm keystroke was in the latter half of the study.

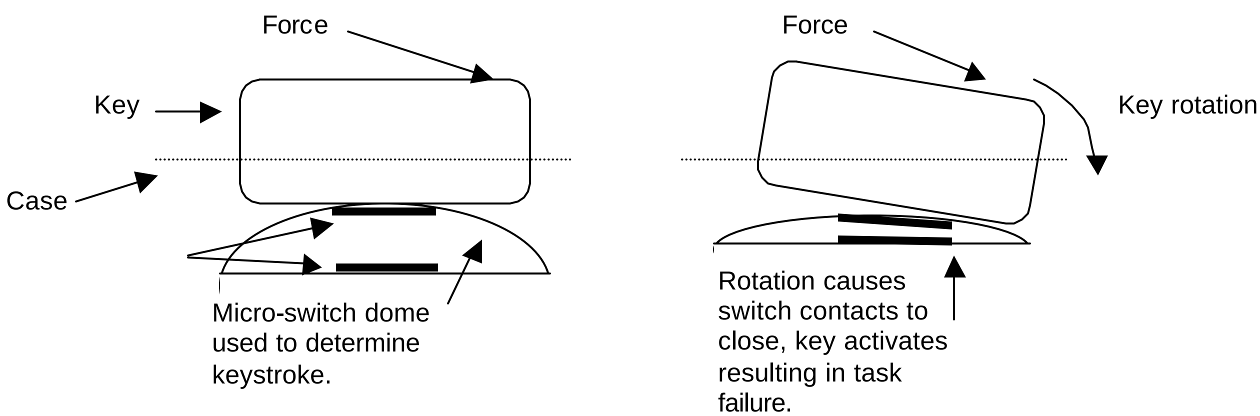


Figure 10.2.3 – Illustration of Possible Cause of Task Failures Due To Lateral Key Movement

Participants with total blindness were unable to differentiate between each level of keystroke. This was particularly surprising, since it was assumed that those with total blindness would have the greatest tactile discrimination ability among the participants.

The 0.5 mm and 0.7 mm levels of keystroke were preferred the most among all groups, while the 0.3 mm level was consistently preferred the least. Differences in preference between the 0.5 mm and 0.7 mm levels were not significant, so a design decision for the 0.5 mm or the 0.7 mm should be based on factors other than preference, such as error tolerance, or the cost and availability of the micro-switch domes that are used to produce a specific keystroke.

The keystroke with the highest ease of use and accessibility ratings differed, depending on the disability group. Participants in the visual and no disabilities groups rated the 0.7 mm level the highest for both ease of use and accessibility, while participants in the physical disabilities groups rated the 0.5 mm level the highest for ease of use, and the 0.3 mm level the highest for accessibility. These differences in ratings slightly nullified one another for the overall measures. Ease of use and accessibility ratings for the 0.7 mm level were slightly higher than those for the 0.3 mm and 0.5 mm levels. The differences in each level of keystroke were so small (0.2 mm), that it is unclear whether participants were able to differentiate between keystroke levels when models were not compared side to side.

The decision to select a level of keystroke should be based on a composite of the objective measures, such as task completion time and the number of task failures, as well as the subjective measures, such as preference, ease of use, and accessibility (Bailey, 1993). Since the subjective measures were relatively similar, as were the task completion times, the decision on which keystroke level to design should be based on the number of task failures. The number of task failures was the lowest with the 0.3 mm level of keystroke, followed closely by the 0.5 mm level. Task completion times and subjective ratings were the best for the 0.7 mm keystroke. Due to the relatively high number of task failures for the 0.7 mm level, and the minimal differences in task completion time and subjective ratings between levels, a design decision should be made for either the 0.3 mm or 0.5 mm levels of keystroke (Figure 10.2.4).

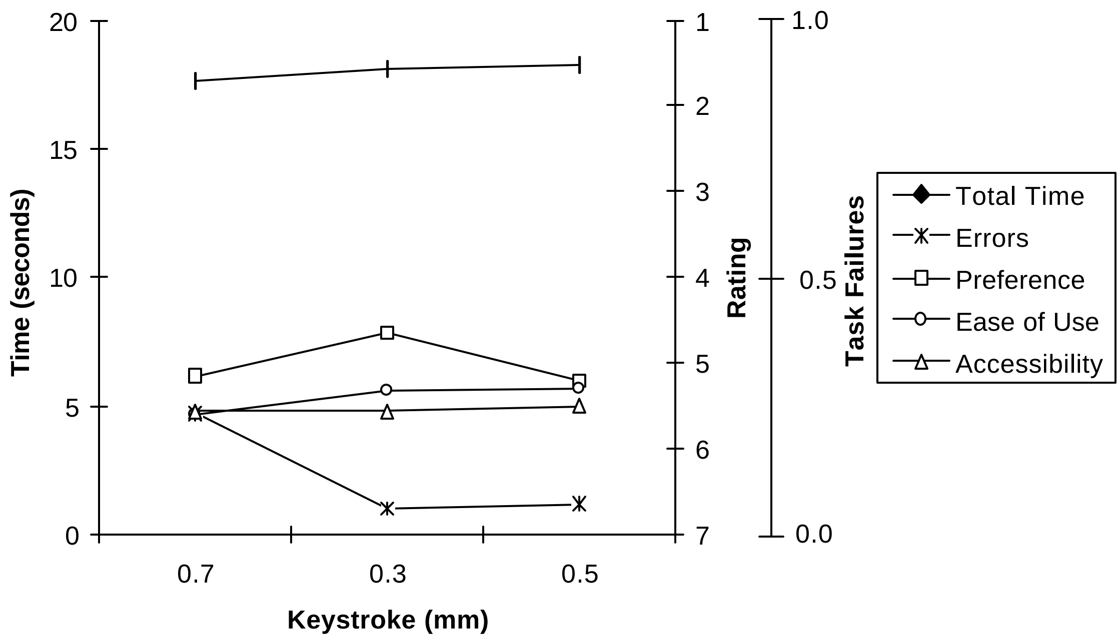


Figure 10.2.4 – Mean Overall Objective and Subjective Measures by Keystroke
(Y-axes are oriented so that values closer to the origin indicate better results.)

Key Height

Task completion times, overall, were the fastest with the 0.7 mm key height. The 0.7 mm key height may have provided the best tactile discrimination, allowing users to find the keys more quickly, with higher confidence that they chose the correct key. Tactile discrimination is especially important for those with low vision, since tactile identification allows users a method to identify keys without the use of sight (Sanders and McCormick, 1993).

Stored Memory Recall, and 911 dialing task completion times, however, were the fastest for participants with visual disabilities when using the 0.5 mm key height. The shorter task completion times when using the 0.5 mm key height may have been due to participant cuing, and experimental error, rather than key height.

The Stored Memory Recall and 911 dialing tasks only required the participants to press two or four keys (the number and the <Send> key). The experimenter read the number before participants completed the task. Tasks were timed from when the first tone was heard (indicating a key was pressed) until the last tone was heard (the <Send> key was pressed). Participants may have reduced the effort required to navigate the keypad by positioning their fingers on, or close to the appropriate keys before beginning the task (cueing), thus decreasing the time to complete tasks (Wickens, 2000). Since only a couple of keys were pressed, cueing, and variations in timing may have affected the overall result. Although each Stored Memory Recall task consisted of three trials, using three different numbers between two and nine, the 911-dialing task repeated the same number each time. After the participant dialed the number once, they were prepared to dial it a second and third time. In some instances, participants dialed the three 911 tasks in quick succession, and in some instances more quickly than the experimenter could time reliably. At least one participant with visual disabilities was asked to pause for several seconds between tasks so that timing was easier. Although issues with timing and participant cueing may have occurred consistently throughout the experiment, the effects of such may have been magnified during these specific tasks.

The number of task failures was the lowest with the 0.3 mm key height. Since the 0.3 mm key height is nearly flush with the case, participants may have used more caution in dialing these models. A flush mounted key provides less tactile feedback, requiring extra time for the user to discriminate the key from the case (Sanders and McCormick, 1993). The extra time required may also be due to the Speed / Accuracy tradeoff, where the user slows performance to ensure they are pressing the proper key (Sanders and McCormick, 1993). The task completion time support this hypothesis, since task completion times were the slowest with the 0.3 mm key height.

The key height that was preferred the most differed by disability group. All participants with visual impairments preferred the 0.7 mm level the most, while those in the physical, and no disabilities groups preferred the 0.5 mm level the most. All groups preferred the 0.3 mm level the least. The preference for either the 0.5 mm or the 0.7 mm levels of key height was significantly higher than that of the 0.3 mm level overall, since the 0.5 mm and 0.7 mm levels provide better tactile information than the 0.3 mm level (Sanders and McCormick, 1993). Differences in preference between the 0.5 mm and 0.7 mm levels were not significant, however.

The groups with visual, and no disabilities rated the 0.7 mm key height the highest for ease of use, while those with physical disabilities rated the 0.5 mm level the highest. All participants rated the 0.3 mm key height the lowest, and these ratings were significantly different than ratings for the 0.5 mm and 0.7 mm levels of key height.

All groups rated the 0.7 mm key height the highest for accessibility, followed by the 0.5 mm and 0.3 mm levels, respectively. Accessibility ratings for the 0.7 mm and 0.5 mm were significantly different than ratings for the 0.3 mm, but not significantly different from each other.

The 0.7 mm and 0.5 mm levels of key height had the fastest task completion times, respectively, and received the highest preference rankings, ease of use, and accessibility ratings overall. The number of task failures was the lowest when using the 0.3 mm level, and highest when using the 0.5 mm level. The 0.7 mm key height was the best in all measures, except for the number of task failures, where the 0.3 mm key height was the best. Differences between the 0.7 mm and 0.5 mm key heights were not significant across any of the measures, so the decision of which key height to produce should be made based on subjective measures, or other factors, such as cost, availability, etc.

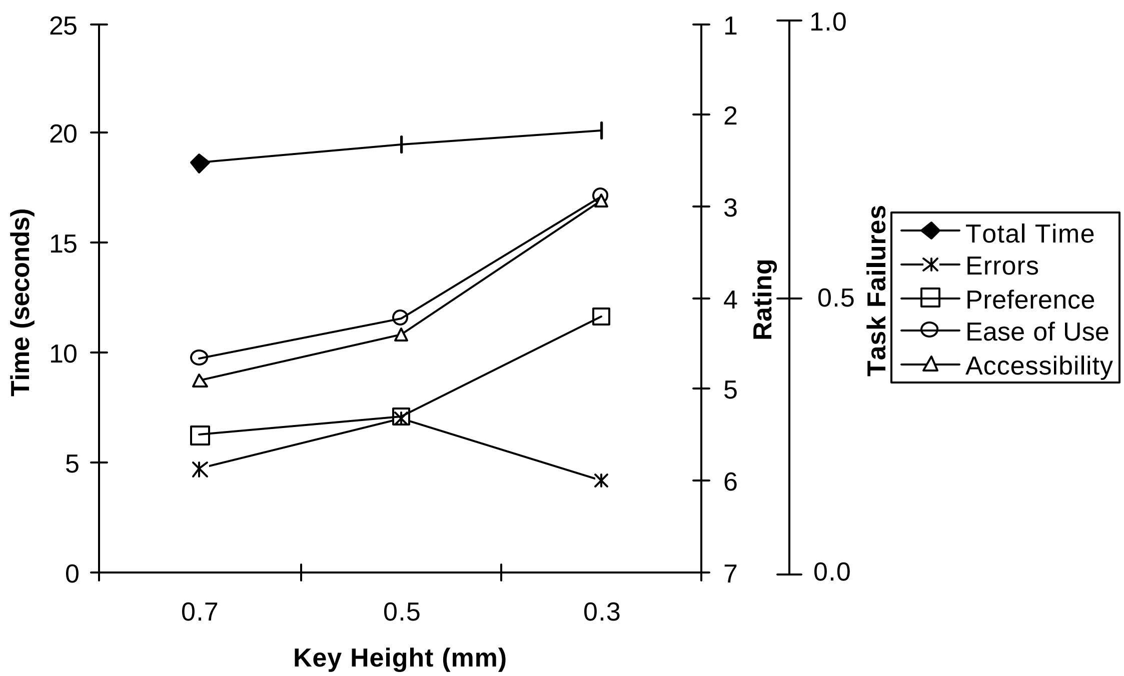


Figure 10.2.5 – Mean Overall Objective and Subjective Measures by Key Height
(Y-axes are oriented so that values closer to the origin indicate better results.)

Position of <Power>, <Send> and <End> Keys

Preference rankings for the <Power> key position were the highest for models with the <Power> key located in a distinct position, separate from the numeric keypad (Table 8.6.1, models 2, 4, and 6). Ease of use and accessibility ratings, however, did not follow this trend closely.

When participants were asked to rank order each telephone based “solely on the position of the <Power> key”, participants held the model, compared the <Power> key position to other keys, and then ranked the models based on which location they preferred. When participants were asked to rate each <Power> key on ease of use, however, participants tried to use the <Power> key to turn each model on. Participants in the MU and ND groups tried this by holding the telephone with one hand and using the thumb to activate the switch. This was difficult to accomplish with model 4 without dropping it, and the ability to reach the <Power> key with the thumb on model 6 was limited. The <Power> key on model 2 proved very difficult to activate even with two hands, since it was very small and mounted flush with the case. Participants without visual disabilities often used a fingernail to activate the switch, while those with visual disabilities had an extremely difficult time finding it, as well as activating it (Figure 10.2.6).

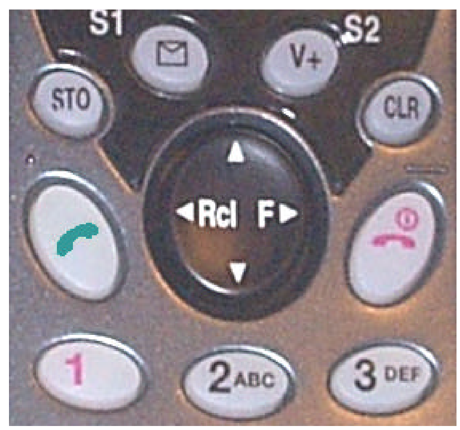
Figure 10.2.6 - <Power> Key Location on Model 2
(located on top right edge of telephone). Photograph enlarged to show detail.

Accessibility ratings followed a similar trend to the ease of use ratings, except for those with legal blindness. Participants with legal blindness rated the accessibility of the <Power> key on model 2 more highly than the ease of use. These participants may have had a strict interpretation of accessibility as “the ability to reach something”, rather than “the ability to reach something and use it” definition that other participants may have used. Boredom, fatigue, and apathy may also account for the conflicting ratings for these participants, since the rating portion was near the end of the experimental session.

Overall preference rankings, as well as ease of use and accessibility ratings for <Send> and <End> key position, were the highest for models with distinct, raised keys that were separate from the numeric keypad, different in shape, and easily identifiable. Models 6 and 8 generally received the highest ratings, while models 4 and 5 were generally rated the lowest.

The consistently high ratings for models 6 and 8 may be due to key size, shape, location, and visibility. Keys on both of these models were larger, and shaped differently than the surrounding keys. The keys on these models were slightly separated from the numeric keypad, and were marked with bright, descriptive icons (Figure 10.2.7). Sanders and McCormick (1993) suggests that users perform the best with controls (keys) with features such as these.

Model 6



Model 8

Figure 10.2.7 - <Send> / <End> Key Configuration on Models 6 and 8
(Note different size and shape, as well as separation from other keys.)

Models 4 and 5 were rated the lowest by several of the participant groups. Model 4 has <Send> and <End> keys that are a different size than the numeric keys, and are located away from the numeric keypad. The function of both keys is not consistent with what users expect, however. The <Send> key is commonly located on the left, and the <End> key on the right. Model 4 has the <End> key on the left, and the <Send> key on the right. The <Send> and <End> keys on model 5 are located in the lower right hand corner, in a stacked configuration. These keys are close to the numeric keypad, relatively small, and the same size and shape as other function keys on the same row. Participants experienced difficulty finding these keys, and difficulty operating those keys with one hand without dropping the telephone (Figure 10.2.8). Since the arrangement of these keys was not consistent with most cellular telephones, user preference may have been reduced (Hix and Hartson, 1993; Nielsen, 1993; Norman, 1988; Schneiderman, 1998; Wickens, 2000).

Model 4 - Located on top left and right

Model 5 - Located in lower right hand corner

Figure 10.2.8 - <Send> / <End> Key Configurations on Models 4 and 5

Display Font Size

Nearly all of the participants rated the 22-point and 36-point display fonts more highly than the 16-point font for preference, ease of use, and accessibility. Participants did not show a strong preference for the normal versus bold weight fonts. Participants believed the 22-point and 36-point display fonts were easier to see than the 16-point font. Those with no disabilities mentioned that they may prefer the 16-point font if the telephone has special capabilities, such as text messaging, email, and Internet access, however.

10.3 - Universal Design Principles

Universal design is the design of products and environments to be usable by the greatest pool of people as is practical, without the need for adaptations or specialized design (Center for Universal Design, 1997). The inclusion of universal design in products can increase market share, as well as increase compliance with regulations such as the Americans with Disabilities Act, Section 255 of the Telecommunications Act, and Section 508 of the Rehabilitation Act (Mondak 2000; Mueller, 1995).

The most common issues observed during the experiment were noted and compared with the universal design principles (Table 2.5.1) to determine how universal design can be applied to cellular telephones to improve accessibility. Each principle is listed below, along with explanations of how each level of independent variable demonstrated, or violated these principles.

Equitable Use

The principle of equitable use refers to a design that is useful and marketable to people with diverse abilities. A design for equitable use allows for consistent operation for all users, avoids segregating or stigmatizing users, and makes the design appealing for all (Center for Universal Design, 1997).

Each level of lateral pitch used during the experiment was relatively equitable in use. Although the participants could use each level of lateral pitch, the 12 mm and 13 mm levels were the most appealing. These levels of lateral pitch were easy to handle, had the lowest task completion times and number of task failures, and had the highest subjective measures such as preference, ease of use, and accessibility.

The 0.3 mm and 0.5 mm levels of keystroke were the most equitable in use. Participants with all ability levels could use these levels of keystroke with little difficulty. The 0.7 mm keystroke seemed to have the least equitable use, since participants with physical disabilities committed the most task failures with this level.

The 0.7 mm key height was the most equitable in use, while the 0.3 mm key height was the least. The 0.7 mm key height allowed all users to find the keys quickly, and confidently. The 0.3 mm key height was relatively flush with the face of the telephone, and users completed tasks more slowly, albeit with fewer task failures when using this level of key height. Although the number of task failures with the 0.3 mm key height was the lowest, so were the ratings for preference, ease of use, and accessibility, indicating that this key height is not appealing to users.

The <Power>, <Send> and <End> keys on models 6 and 8 were the most equitable in use. These keys were easy to identify, prominent, separate from the numeric keypad, and could be easily pressed during one-handed or two-handed operation. Participants also rated models 6 and 8 the highest for <Power>, <Send>, and <End> key ease of use and accessibility, indicating designs that are appealing to users.

The design of the <Power> key on models 2 and 5 (see Table 8.6.1) were not the most equitable in use for those with visual disabilities or limited dexterity. These key designs, especially on model 2, were small and difficult to press. The <Power> key on model 2 was difficult to see, and some users had trouble finding it due to the size and location.

The 22-point and 36-point display fonts were the most equitable in use, since those with legal blindness were able to see them clearly. The 16-point font was too small to effectively use for these participants, and was the least appealing to all participants.

Flexibility In Use

Flexibility In Use refers to a design that accommodates a wide range of individual preferences and abilities. Flexibility in use allows users a choice in the method they wish to use a device, accommodates use with either hand, or facilitates accuracy and precision (Center for Universal Design, 1997).

The 12 mm and 13 mm models (models 8 and 6) were the most flexible in use, since users could easily activate the <Power> key, dial, and use the <Send> and <End> key on these models using one or two hands. One-handed dialing with the 10 mm and 11 mm models (models 9 and 1, respectively) was difficult, since these models were difficult to grip due to their width. Models 2, 4 and 5 (see Table 8.6.1), also had poor flexibility in use, since users were required to use both hands to turn them on, or risked dropping the telephone if operated with one hand.

The 0.5 mm keystroke, and 0.3 mm key height facilitated the greatest accuracy and precision since the number of task failures were the lowest while using these levels. The accuracy and precision required when using the 0.5 mm keystroke and 0.3 mm key height, however, come at the cost of slower task completion time, and lower subjective evaluation scores (preference, ease of use, and accessibility).

The 22-point font provides the greatest flexibility in use. This font is large enough to be seen easily, yet is small enough to be used with telephones with other functions, such as text messaging and email. The 16-point font was the least flexible in use, since users with legal blindness had trouble seeing it and may read the numbers incorrectly.

Simple and Intuitive Use

The principle of Simple and Intuitive Use refers to a design that is easy to use and understand, regardless of experience, knowledge, language, etc. Designs that incorporate simple and intuitive use eliminate unnecessary complexity, are consistent with user expectations, and provide effective prompting and feedback (Center for Universal Design, 1997).

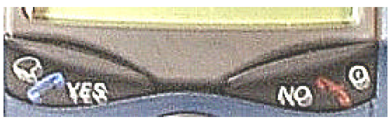
According to Mueller (1995), simple and intuitive use is especially important for casual users. Users interviewed in the Smith-Jackson et al. (2001) study indicated that when they used a cellular telephone, it was only intermittently, and for very short periods of time. Since the cellular telephone was used intermittently, simple and intuitive use is especially important if users are to complete calls quickly and efficiently.

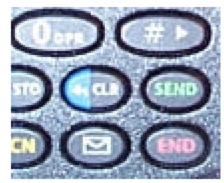
All levels of lateral pitch, keystroke, and display font were relatively simple and intuitive to use. The 0.3 mm key height, however, was not as simple and intuitive to use as the other levels of key height. Participants with visual disabilities required extra time to complete tasks since the keys were difficult to discern from the face of the telephone.

The <Power>, <Send> and <End> keys on models 6 and 8 were the most simple and intuitive to use, due to the location, size, use of colors, and descriptive icons denoting their function. The <Power> keys on these models were placed in positions that allowed for easy operation with one or both hands. The buttons were large and easy to press, and were labeled with a commonly used "Power" symbol. The <Send> and <End> keys on these models were also large, labeled with descriptive icons, and easy to operate with one or both hands.

The <Power> keys on models 2 and 3 (see Table 8.6.1) were not as simple and intuitive to use as the <Power> keys on other models. The <Power> key on model 2, located on the top of the telephone, is very difficult to find, and even more difficult to activate. Users with visual disabilities or limited dexterity may not be able to operate the <Power> key on this model. The <Power> key on model 3 is labeled "No", which may cause confusion.

The <Send> and <End> keys on models 3, 4, and 5 were the least simple and intuitive to use. The <Send> and <End> keys on model 3 were labeled “Yes” and “No”, instead of “Send” and “End”. The <Send> and <End> keys on model 4 were small, and placed in the reverse order of the <Send> and <End> keys on other cellular telephone models. The <Send> and <End> keys on model 5 were in a stacked configuration on the lower right hand corner of the telephone, making one-handed operation difficult. Although designs such as these are novel, it may be best to continue to design models with the <Send> and <End> keys located near the top of the telephone, with the <Send> key on the left, and the <End> key on the right (Norman, 1988).





Model 3 Model 4 Model 5

Figure 10.3.1 - <Send> / <End> Keys on Models 3, 4, and 5

Perceptible Information

The principle of Perceptible Information communicates necessary information to the user, regardless of ambient conditions or sensory abilities. Perceptible information uses redundant modes of information presentation, adequate information contrast to maximize the legibility of essential information, and compatibility with techniques used by people with sensory limitations (Center for Universal Design, 1997).

The 0.7 mm key height was the most perceptible, since this height was easy to discriminate from the face of the telephone. The 0.3 mm key height was the least perceptible to users with visual disabilities. This key height was relatively flush with the face of the telephone, which made it difficult to discern the keys from the telephone face.

The 22-point and 36-point fonts were much more perceptible than the 16-point font, since they were larger and much easier to see. The 16-point display font was the least perceptible to all participants.

The <Power>, <Send>, and <End> keys on models 6 and 8 were the most perceptible. These keys were larger, shaped differently, and located away from surrounding keys. Icons to denote the function of these keys were also large and descriptive. The <Power>, <Send>, and <End> keys on models 1, 5, 7, and 9 had limited perceptibility. The similarity in size and shape of these keys to the numeric keys, as well as the lack of separation from the numeric keypad, reduced the ability to differentiate these keys from the numeric keys on these models (Figure 10.2.1). Keys that have different functions should be shaped differently and located away from other keys so users can differentiate them easily (Sanders and McCormick, 1993).

Tolerance for Error

Designs that incorporate Tolerance for Error minimize hazards and adverse consequences of accidental actions, provide fail-safe features and warn of errors, and discourage unconscious actions (Center for Universal Design, 1997). The 12 mm and 13 mm levels of lateral pitch were the most error tolerant. Due to the spacing between keys, extreme accuracy and fine motor skills were not required to avoid errors while dialing. The lateral pitch of the 10 mm model was the least tolerant for errors, due to the lack of space between keys and the extra precision required to press the correct key.

The 0.5 mm keystroke was the most tolerant for error as evidenced by the low number of task failures. The 0.7 mm keystroke was the least tolerant for error, since keys with this keystroke were often inadvertently activated due to lateral movement.

The 0.3 mm and 0.7 mm key heights were the most error tolerant, since the number of task failures with these keys was the lowest. Although the 0.3 mm key height was the most error tolerant, participants completed tasks more slowly with this key height, and rated this key height poorly for preference, ease of use, and accessibility.

Models 6 and 8 had the greatest tolerance for error in the design of the <Power>, <Send>, and <End> keys. These keys were large, prominent, clearly labeled, and located away from the numeric keypad. The design of the <Power> keys on models 2 and 3 had poor tolerance for error. The <Power> key on model 2 was difficult to find and activate, while the <Power> key on model 3 was labeled “No”. The <Send> and <End> keys on models 5 and 7 were the same size and shape as the numeric keys, and the location of these keys on model 5 made one-handed use without dropping the telephone difficult.

Low Physical Effort

Designs that incorporate the principle of low physical effort can be used efficiently, and comfortably with a minimum of fatigue. These designs allow users to maintain neutral postures, use reasonable operating forces, and minimize repetitive actions and physical effort (Center for Universal Design, 1997).

Models 1, 6, and 8 (11 mm, 13 mm, and 12 mm lateral pitches, respectively) could be used with low physical effort, however model 9 (10 mm lateral pitch) could not. Due to the width thickness, and surface finish of model 9, holding and maintaining a grip on the telephone may have been difficult.

All levels of keystroke, and the 0.5 mm and 0.7 mm levels of key height could be used with low physical effort. The 0.3 mm key height may have required more physical effort than the other key heights, since it was difficult to differentiate from the face of the telephone.

The <Power>, <Send>, and <End> keys on models 1, 6, and 8 required the least physical effort, since they were easy to find by touch. The <Power> key on model 2, as well as the <Send>, and <End> keys violated the principle of low physical effort. The <Power> key on model 2 was small, difficult to find, and had to be pressed and held with a fingernail in order to activate it. The <Send> and <End> keys on model 5 are located in the lower right hand corner of the telephone. Greater physical effort may have been required to operate the <Send> and <End> keys on this model than was required on the other models.

Size and Space for Approach and Use

Designs that incorporate the principle of size and space for approach and use provide for approach, reach, manipulation, and use regardless of a user's size, posture, or mobility. Additionally, these designs provide a clear line of sight to important elements for seated or standing users, make reach comfortable for users, and accommodate different hand or grip sizes (Center for Universal Design, 1997).

The 12 mm and 13 mm levels of lateral pitch provided appropriate size and space for approach and use. These levels of lateral pitch were large enough to allow users with different hand sizes, and physical ability levels the ability to use the keypad confidently. The 11 mm lateral pitch may be acceptable for some users, and too small for others. The 10 mm lateral pitch was generally too small for participants to use comfortably.

All levels of keystroke, and key height provided for size and space for approach and use. Users were able to approach, reach, and manipulate each level of keystroke and key height relatively well.

The <Power>, <Send>, and <End> keys on models 1, 6, and 8 provided the most appropriate size and approach for use. The placement of these keys allowed users to activate them easily, using one or both hands, with little risk of dropping the telephone. The <Power> key on models 2 and 5 did not provide adequate size and space for approach and use. The <Power> key on model 2 was very small and difficult to use, while the <Power>, <Send>, and <End> keys on model 5 was difficult to use with one hand comfortably, and without the risk of dropping the telephone.

The inclusion of the universal design features outlined above should assist cellular telephone manufacturers in the design of telephones that are more usable and accessible to those with disabilities, as well as those without. Products that are more usable and accessible may allow manufacturers to enjoy a greater share of the cellular telephone market (Mueller, 1995). Additionally, telephones that incorporate these features can comply with important disability-related regulations, such as the Americans with Disabilities Act, Section 508 of the Rehabilitation Act, and Section 255 of the Telecommunications Act (Mondak, 2000).

10.4 - Usability

Usability principles include learnability, memorability, efficiency, errors, and satisfaction (Hix and Hartson, 1993; Jordan, 1998; Nielsen, 1993; Preece, 1993). The most common usability issues observed were analyzed to determine which usability principles were being violated. Solutions to these, and the issues identified in the Universal Design section (section 10.3) will be addressed through the use of product-specific guidelines in section 10.5.

Learnability

Each level of lateral pitch, keystroke, and key height was relatively easy to learn. Since users had previous experience with using a telephone, very little learning was required.

The design of special function keys on models 1 and 9 (10 mm and 11 mm, respectively) used in the dialing tasks, and the <Send> and <End> keys on models 3, 5, and 7, may present a problem with learnability. These keys are relatively similar in size, shape, and location to the numeric keys. These similarities require users to learn and memorize the specific placement of keys, rather than relying on tactile information for identification. Learning and memorization of these locations may be troublesome, especially if the telephone is not used frequently (Nielsen, 1993).

The <Send> and <End> keys on model 4 may present a problem with learnability. Although these keys are separate and easy to distinguish from the numeric keys, the function of each is the opposite of that commonly used on cellular telephones. The <Send> key is commonly located on the left, and the <End> key located on the right. The <Send> key on model 4, however, is located on the right, and the <End> key on the left. This may not be a learnability issue for new cellular telephone users, though it may present an issue for users with previous cellular telephone experience. The arrangement of function keys such as this should remain consistent with arrangements commonly used with the <Send> key on the left, <End> key on the right (Norman, 1988).

Memorability

The levels of lateral pitch and keystroke did not present an issue with memorability. The 0.3 mm level of key height may present an issue with memorability, however. The 0.3 mm key height was relatively flush with the face of the telephone. Due to the lack of protrusion, tactile identification of specific keys becomes more difficult. Since the keys did not protrude very much, users must memorize where specific keys are if they wish to dial the telephone by touch.

The <Send> and <End> keys on model 4 may present an issue with memorability. The reversed function of these keys requires casual users to remember this new configuration, especially if they dial by touch and not by sight.

Efficiency

The 12 mm and 13 mm levels of lateral pitch were relatively efficient, as were the 0.3 mm and 0.5 mm levels of keystroke, and the 0.5 mm and 0.7 mm levels of key height. The 10 mm and 11 mm levels of lateral pitch may have reduced efficiency due to the lack of separation between the keys. This lack of separation increases the accuracy required to place the fingers on the proper keys, or else a task failure may result. Since dialing a telephone is a target selection task, there is a relationship between the speed of target acquisition and the accuracy required to acquire that target. More specifically, as accuracy demands are increased, speed is decreased. The lack of separation between the keys on the 10 mm and 11 mm keys increased the accuracy required, thus decreasing the speed in which tasks were completed, resulting in decreased efficiency.

The 0.7 mm keystroke, and the 0.3 mm key height had reduced efficiency due to the increase in task failures with the 0.7 mm keystroke, and the increase in task completion times with the 0.3 mm key height. Participants committed the most task failures when using the 0.7 mm keystroke. Once a task failure was committed, the user had to fix the error that caused the task failure in order to complete the dialing task successfully. This increase in time required to correct these errors reduced efficiency. The 0.3 mm key height was difficult to differentiate from the face of the telephone. Due to this, users with visual disabilities or those that wish to dial by touch are required to dial more slowly to ensure the proper key is selected, thus reducing efficiency.

The 22-point and 36-point display fonts may be more efficient than the 16-point font, especially for users with visual disabilities. The 22-point and 36-point fonts are larger than the 16-point font, and relatively easier to see. Users with legal blindness may have difficulty recognizing characters displayed in the 16-point font. The difficulty in recognizing which characters are displayed can reduce efficiency by slowing the user, or facilitating a dialing error that must be corrected.

The <Power>, <Send>, and <End> keys on models 6 and 8 were relatively efficient, while the <Power> key on model 2, and the <Power>, <Send>, and <End> keys on models 4 and 5 were relatively inefficient. The design of these keys allowed users to easily identify and use the <Power>, <Send>, and <End> keys on models 6 and 8 easily, thus increasing efficiency (Sanders and McCormick, 1993). The <Power> key on model 2 was relatively difficult to find, as well as use. The extra time required to activate this key reduced the efficiency. The <Power>, <Send>, and <End> keys on models 4 and 5 were relatively inefficient, due to placement and inconsistency (Norman, 1988). The <Power> keys on both models were located in the lower left hand corner of the telephone. Extra care was needed to activate this key using one hand, or the user risked dropping the telephone. The <Send> and <End> keys on model 5 were located in the lower right hand corner, again, requiring extra care to avoid dropping the telephone when operated with one hand. The <Send> and <End> keys on model 4 were placed in the reverse order of most cellular telephones, which may reduce efficiency due to the need to correct errors committed by users expecting an arrangement similar to other telephones.

Errors

The prevention of errors (error tolerance) and the ability to easily undo errors is an important principle for both usability, as well as universal design. Refer to “Tolerance for Error” in section 10.3 (Universal Design) for a discussion of which attributes provided the greatest and least error tolerance.

Satisfaction

The 12 mm and 13 mm levels of lateral pitch were the most satisfying to use. User satisfaction with these levels is apparent due to the relatively consistent preference, ease of use, and accessibility ratings. The 12 mm model was rated the highest for these measures, followed closely by the 13 mm model. Participants were the least satisfied with the 10 mm level of lateral pitch. This dissatisfaction may have been due to the lack of separation between keys, or issues such as the flush keys or the slick finish of the model used.

The 0.7 mm level of keystroke was the most satisfying to use, as indicated in the preference, ease of use, and accessibility ratings. Due to the minute differences between keystroke levels (0.2 mm), participants may not have been able to tell the difference between each level. The higher level of satisfaction with the 0.7 mm keystroke may be due to chance, and not due to users feeling less satisfied with the other levels.

The 0.7 mm key height was the most satisfying to use, as indicated in the preference, ease of use, and accessibility ratings. The 0.3 mm key height provided the least satisfaction. The 0.7 mm key height was easy for participants to differentiate from the case by touch due to the protrusion. The 0.3 mm key height, which was relatively flush with the case, was difficult to differentiate from the case by touch. Participants were more satisfied using a key they could easily distinguish from the body of the telephone, as opposed to one that was difficult to distinguish.

The 22-point and 36-point fonts were the most satisfying to participants. These fonts were larger, and much easier to see than the 16-point fonts. Participants may have been more satisfied with the 16-point font if the telephone had special features such as text messaging, or email and Internet access.

The <Power>, <Send>, and <End> keys on models 6 and 8 were rated the highest for satisfaction, as indicated by the preference, ease of use, and accessibility ratings. These keys were large, prominent, clearly labeled, and separate from the numeric keypad, allowing for easy one-handed or two-handed operation.

The <Power> keys on models 2, 3, and 5 were the least satisfying to use. The small size, poor labeling, and difficult to reach locations, made use difficult, reducing user satisfaction. Size, ability to reach the keys, and proper labeling are important features that increase the ease in which a design can be used, which has an effect on satisfaction (Sanders and McCormick, 1993; Nielsen, 1993).

The <Send> and <End> keys on models 4 and 5 were the least satisfying to use. These keys were in locations that were difficult to reach when using the telephone with one hand without the risk of dropping the telephone. These difficulties lead to low satisfaction as indicated by the preference, ease of use, and accessibility ratings.

10.5 - Design Guidelines

Product-specific guidelines were created based on objective and subjective results made during the study observations. Guidelines are listed by the issue leading up to the guideline, the universal design and usability principles the guideline complies with, the rationale behind the guideline, and an example of good design (Table 10.5.1).

Designers should not view these guidelines as a comprehensive list or final solution to prevent all usability issues. These guidelines should be used in conjunction with other information not available in this study, and reviewed judiciously along with other considerations to determine when the application of these guidelines is appropriate. Since usability is an iterative process, the implementation of these guidelines should be subject to further testing to ensure that usability was improved. The appropriate incorporation of these guidelines should help to increase the usability of cellular telephones for users with disabilities, as well as those without.

Table 10.5.1 - Universal Design and Usability Guidelines to Improve Cellular Telephone Usability and Accessibility

Issue:	Universal Design Principle:	Usability Principle:	Guideline:	Rationale:	Design Example:
Amount of lateral pitch to optimize objective and subjective measures.	Equitable use Flexibility in use Simple and intuitive use Tolerance for error	Efficiency Errors Efficiency	Lateral pitch should be greater than 10 mm, but less than or equal to 13 mm.	Objective and subjective measures were the best as the lateral pitch increased, up to the 12 mm level, and then decreased at the 13 mm level.	Model 8 (Table 8.6.1).
Lateral separation between columns of keys.	Equitable use Flexible use Simple and intuitive use Perceptible information Tolerance for error Low physical effort	Learnability Efficiency Errors Satisfaction	Lateral separation between key columns should be at least 2.5 mm, but less than or equal to 4.5 mm.	Objective and subjective measures increased as this distance increased, with peak measures occurring at 3.4 mm. Objective and subjective measures declined beyond 3.4 mm.	Model 1, 6, 8 (Table 8.6.1).
Vertical separation between numeric keypad rows.	Equitable use Flexibility in use Simple and intuitive use Perceptible information	Learnability Efficiency Errors Satisfaction	Vertical separation should be at least 2.0 mm, but not greater than 2.5 mm.	Objective and subjective measures were the increased as vertical separation increased, with the best objective and subjective measures occurring at 2.3 mm. Objective and subjective measures decreased beyond 2.3 mm.	Model 1, 6, 8 (Table 5.6.1).

	Tolerance for error Low physical effort				
Difficulty identifying <Send>, <End>, and special function keys due to proximity to numeric keypad	Equitable use Flexibility in use Simple and intuitive use Perceptible information Tolerance for error Low physical effort	Learnability Memorability Efficiency Errors Satisfaction	The <Send>, <End>, and special function keys should be separated from the numeric keypad by 150% of the distance used to separate numeric keypad rows.	The increased spacing between the numeric keypad row and the <Send>, <End>, and special function key rows will allow users to more easily identify keys that are different from the numeric keypad keys by touch.	Models 6 and 8 (Table 8.6.1).
Location, separation, shape, and orientation of special function keys.	Equitable use Flexibility in use Simple and intuitive use Perceptible information Tolerance for error Low physical effort	Learnability Memorability Efficiency Errors Satisfaction	Special function keys should not be located between the top row of the numeric keypad and the <Send> and <End> keys, or immediately above the <Send> and <End> keys, unless there is sufficient space between key rows, or keys are substantially different in size, orientation, or shape.	Special function keys that are located very close to the numeric keypad or the <Send> and <End> keys were difficult to differentiate from other keys. Extra separation, or differences in shape, orientation, and size will allow users to differentiate these keys from others by touch, resulting in inadvertent activation.	Models 6 and 8 (Table 8.6.1).
Landmark references difficult to identify by touch.	Equitable use Flexibility in use	Learnability Memorability Efficiency	Reference landmarks should be placed on the keys instead of on the telephone body.	Reference landmarks placed on the telephone keys are easier to identify than those placed on the telephone body. Landmarks placed on the body can be difficult to feel, depending on the	Braille dot on <5> and bar on <Send> and <End> keys.

	Simple and intuitive use Perceptible information Tolerance for error Low physical effort	Errors Satisfaction		key height.	
Reduce lateral key flexion to a minimum.	Equitable use Flexibility in use Tolerance for error Low physical effort	Efficiency Errors Satisfaction	Lateral key flexion should be kept to a minimum. Lateral key flexion can be reduced by using a keystroke that is less than 0.7 mm.	Lateral key flexion was a possible contributor to task failures for those with physical disabilities. Keys were inadvertently activated when brushed against at an angle, leading to errors.	0.3 mm and 0.5 mm keystroke levels.
Use a key height that allows for fast task completion times, low task failures, and high satisfaction.	Equitable use Flexibility in use Simple and intuitive use Tolerance for error Low physical effort	Efficiency Errors Satisfaction	Key heights should be greater than 0.3 mm, but less than or equal to 0.7 mm.	Task completion times and subjective measures were the lowest when using the 0.3 mm key height, and highest when using the 0.7 mm key height. Task failures, however, increased when the 0.7 mm height was used.	0.5 mm and 0.7 mm key heights used on Model 8 (Table 8.6.1).
One-handed operation without risk of dropping telephone.	Equitable use Flexibility in use Simple and intuitive use	Efficiency Errors Satisfaction	Numeric keys, as well as the <Power>, <Send> and <End> keys should be easy to operate with one hand, with minimal risk of dropping the telephone.	Depending on the placement of specific keys, one-handed operation was difficult without dropping the telephone. Dropping the telephone leads to task failures and reduced user satisfaction.	All models used, except for Model 5 (Table 8.6.1).

	Tolerance for error Low physical effort				
Ensure display fonts are easy to see.	Equitable use Flexibility in use Tolerance for error	Efficiency Errors Satisfaction	Display font sizes should be between 22-point and 36-point font for normal telephone functions (dialing numbers and not email, text messaging, or Internet access).	The 22-point and 36-point fonts allow for easy viewing for users with legal blindness and normal vision.	22-point and 36-point fonts used on Model 1 during the experiment.
Function of keys difficult to see or understand.	Equitable use Flexibility in use Simple and intuitive use Tolerance for error Low physical effort	Learnability Memorability Efficiency Errors Satisfaction	All keys should have clear, descriptive icons and labels. Icons and labels should be tested with users to ensure users can identify them.	Labels on keys such as the function keys were difficult to understand. This difficulty in understanding may lead to unused functionality, or task failures.	Models 6 and 8 (Table 8.6.1).
Color contrast between keys and labels.	Equitable use Flexibility in use Simple and intuitive use Tolerance for error Low physical effort	Learnability Memorability Efficiency Errors Satisfaction	Color contrast between the key color and the label should allow for easy viewing in low-light conditions.	Key / label combinations with poor contrast may be very difficult to see in low-light conditions.	All models, except for Model 3 (Table 8.6.1).

10.6 - Limitations

Although this study did identify specific keypad attributes that may have contributed to faster task completion time, lower numbers of task failures, and higher subjective evaluation scores, some issues may limit the generalization of these results. Participants in this study were included based on self-reports of disability, and placed in broad disability categories based on those reports. Tests to quantify the severity of visual and physical disabilities were not used. Vision, strength, range of motion, and dexterity tests were considered, but due to the limited availability of participants, and the wide range of severities of specific disability types, testing was not feasible. While this may decrease internal validity, and the ability to generalize for a specific disability, the external validity and the ability to generalize the results in an applied setting may be increased. The general consumer population consists of people with a wide range of visual and physical abilities, and not people in discrete categories (Vanderheiden, 1990).

Participants were required to dial the telephones by holding the telephone with one hand, and dialing with the other. Dialing the telephone in this manner may have had an effect on task completion times and the number of task failures, since it may not have been the dialing method the participant was accustomed to (such as one-handed dialing using the thumb). Due to the length of the experiment, one-handed operation was not feasible, because of fatigue and the inability to keep dialing methods consistent among participants.

Boredom and fatigue may have contributed to the number of task failures that occurred during the experiment. Participants were required to dial a minimum of 256 numbers, and due to the mundane nature of these tasks, may have committed more task failures due to boredom and fatigue, rather than telephone design. Telephone presentation order was partially counterbalanced in the hopes that any errors resulting in task failures, which may be attributed to boredom and fatigue, occurred at random across telephones. Research by Pan, Shell, and Schleifer (1994) indicates that accuracy variability (the difference in task failures) for repetitive data entry tasks is not significant across trial periods, and a one-way ANOVA of presentation order and task failures was not significant. However, transcription errors and errors of omission were removed from the analysis and were not included in the one-way ANOVA. Although these errors were removed from the analyses due to no assignable cause, they did occur. Errors such as this may not occur as frequently when a user is dialing numbers intermittently, as is done with actual use.

Every effort was made to record task completion time for each trial as accurately as possible. Due to the manual nature of recording task completion times, there may be slight errors in the times recorded. The task completion times were measured to the closest $1/10^{th}$ of a second using a stopwatch. Variation in when the stopwatch was started and stopped may contribute to systematic errors. The use of an automated system to record task completion times and errors would help to prevent this in future studies.

10.7 - Future Study Opportunities

The results of this study, as well as feedback from participants in the previous Smith-Jackson et. al (2001) study, may provide opportunities to further study the usability of cellular telephones among those with disabilities. These opportunities include increasing the number of participants, using an improved timing method, testing menu navigation, testing audio displays, and a usability evaluation of models where product-specific guidelines were implemented.

Due to the limited availability of eligible participants in the area, only three participants were included in each disability group (legal blindness, total blindness, etc.). While the results of several small groups can be aggregated to produce an “Overall” result that is important for universal design, as well as marketing efforts, the effects of each attribute on the performance of each disability group had low statistical power. Law and Vanderheiden (1999) suggest methods to simulate specific disabilities that could be used to increase the statistical power, however, they caution that simulation methods should be used only in initial product screening, and not final user testing. Future studies could repeat this experiment using much larger sample populations with real and simulated disabilities, depending on stage of user testing, to increase statistical power.

The timing method used in this study was relatively rudimentary, and subject to systematic error. Future studies could repeat this experiment using an electronic apparatus that can determine task completion times accurately. The use of an apparatus that could catch errors as they occur would be helpful as well.

This study did not address menu navigation. Menu navigation may be important for some users to set up the telephone, and to complete specific tasks. The usability of menus should be tested among users with disabilities in the future. Users with cognitive disabilities should be included in these studies as well.

In the previous study by Smith-Jackson et. al (2001), participants with visual impairments expressed interest in telephones with an audio display, that allowed them to hear what number was being entered. This study could be repeated, comparing the task completion time and number of task failures committed using basic telephone models, and models with audio displays.

The usability process is iterative. It must be repeated to ensure that usability is improved. The results of this study may be used to produce a new cellular telephone prototype using the keypad and display attributes that were optimal. Future study opportunities could focus on usability testing these prototypes, to determine if the attributes selected improved usability, and to determine if other attributes can be improved as well.

11.0 - Cellular Telephone Study Conclusions

The usability of specific cellular telephone attributes was studied among people with visual and upper extremity disabilities. Objective and subjective methods were used to study attributes such as lateral key pitch, keystroke, and key height. Subjective methods were used to study <Power>, <Send>, and <End> key position, as well as display font size. Objective methods included measures of task completion time and the number of task failures, while subjective methods included rank orders of preference, as well as ratings for ease of use and accessibility.

Task completion times were the fastest when using the 12 mm lateral pitch, the 0.7 mm keystroke, and the 0.7 mm key height. The number of task failures was the lowest with the 12 mm lateral pitch, 0.5 mm keystroke, and the 0.3 mm key height. Subjective measures were the best for the 12 mm lateral pitch, the 0.7 mm keystroke, and the 0.7 mm key height. Although task completion times and the subjective measures were the best with the 0.7 mm keystroke, the 0.5 mm keystroke should be used in a design due to the high number of task failures with the 0.7 mm level. Subjective measures were the best for the <Power>, <Send>, and <End> keys on study models 6 and 8, which have large keys located in prominent locations. Subjective measures were the best for the 36-point font.

Each level of lateral pitch, keystroke, key height, display font size, and <Power>, <Send>, <End> key position was evaluated for usability and accessibility using usability and universal design principles. From these principles, and study results, product-specific guidelines were created. These product-specific guidelines can be incorporated in future designs to create cellular telephones that are more usable and accessible to those with visual and upper extremity physical disabilities. Designs that increase accessibility may create a “curb-cut effect”, which results in a product that is more usable for those without disabilities. Cellular telephones that are more usable and accessible may capture greater market share, as well as comply with regulations such as the Americans with Disabilities Act, Section 508 of the Rehabilitation Act, and Section 255 of the Telecommunications Act (Mondak, 2000; Mueller, 1995).

12.0 - References

Access to Telecommunications Equipment and Customer Premises Equipment by Individuals with Disabilities (1997), Telecommunications Access Advisory Committee.

Americans With Disabilities Act of 1990, 28 CFR 35 Title II - Subpart E (1990).

Bailey, R. W. (1993). Performance vs. Preference. Proceedings of the 37th Annual Meeting of the Human Factors and Ergonomics Society, Seattle, WA, HFES.

Bass, L. (2001). Interaction Technologies: Beyond the Desktop. In C. Stephanidis (Ed.), *User Interfaces for All - Concepts, Methods, and Tools* (pp. 81-95). Mahwah, NJ: Lawrence Erlbaum Associates.

Benyon, D., Crerar, A., and Wilkinson, S. (2001). Individual Differences and Inclusive Design. In C. Stephanidis (Ed.), *User Interfaces For All - Concepts, Methods, and Tools* (Vol. 1, pp. 21-46). Mahwah, NJ: Lawrence Erlbaum Associates.

Bikson, T. K., and Panis, C. W. A. (1995). Computers and Connectivity: Current Trends, *Universal Access To Email - Feasibility and Societal Implications* (Vol. 1, pp. 13-40). Santa Monica, CA: RAND.

Campbell, J. L. (1996). The development of human factors design guidelines. *International Journal of Industrial Ergonomics* **18**: 363-371.

Camtasia. (2001). East Lansing, MI, Techsmith.

Carter, J. (1999). Incorporating standards and guidelines in an approach that balances usability concerns for developers and end users. *Interacting with Computers* **12**: 179-206.

Center for Universal Design, N. C. S. U. (1997). *What is Universal Design? - Principles of Universal Design*, [Web page]. The Center For Universal Design - North Carolina State University. Available: http://www.design.ncsu.edu/cud/univ_design/princ_overview.htm [2001, August 8].

College of Engineering Computing Requirements(2001)., [Internet / Web page]. Virginia Polytechnic Institute and State University. Available: http://www.eng.vt.edu/ [2001, August 31].

Edwards, A. D., Pitt, I. J., Brewster, S. A., and Stevens, R. D. (1995). Multiple modalities in adapted interfaces. In A. D. Edwards (Ed.), *Extra-Ordinary Human-Computer Interaction* (pp. 221-243). New York: Cambridge University Press.

Emiliani, P. L. (2001). Special Needs and Enabling Technologies: An Evolving Approach to Accessibility. In S. Constanidis (Ed.), *User Interfaces for All - Concepts, Methods, and Tools* (1 ed., Vol. 1, pp. 97-113). Mahwah, N.J.: Lawrence Erlbaum Associates.

Employment Statistics for People Who Are Blind or Visually Impaired: 1994-1995(2001, February 23)., [Web page]. American Foundation for the Blind. Available: http://www.afb.org/info_document_view.asp?documentid=1529 [2001, September 2].

Hartson, R. H., and Castillo, J. C. (1998). *Remote evaluation for post-deployment usability improvement.* Paper presented at the Working Conference on Advanced Visual Interface (AVI '98), L' Aquila, Italy.

Henninger, S., Haynes, K., Reith, M.W. (1995). *A Framework for Developing Experience-Based Usability Guidelines*. Symposium on Designing Interactive Systems, Ann Arbor, MI, ACM Press.

Hix, D., and Hartson, R. H. (1993). *Developing User Interfaces - Ensuring Usability Through Product and Process*. New York, New York: John Wiley and Sons, Inc.

Hypertext Research v. 2.0.3 (2000). Randolph, MA, Researchware.

Jacobs, S. (1999). *Fueling the Creation of New Electronic Curb Cuts*, [Web page]. The Center For An Accessible Society. Available: http://www.accessiblesociety.org/topics/technology/eleccurbcut.htm [2001, September 8].

Jordan, P. W. (1998). *An Introduction To Usability*. Bristol, PA: Taylor and Francis, Ltd.

Kay, J. (2001). User Modeling for Adaptation. In C. Stephandis (Ed.), *User Interfaces for All - Concepts, Methods, and Tools* (1 ed., Vol. 1, pp. 271-294). Mahwah, N.J.: Lawrence Erlbaum Associates.

Kaye, S. (2000). *Disability and the Digital Divide - Abstract*, [Web page abstract]. Published by U.S. Department of Education, National Institute on Disability and Rehabilitation Research (NIDRR). [2001, August 28].

Langton, A. J., and Ramseur, H. (2001). Enhancing employment outcomes through job accommodation and assistive technology resources and services. *Journal of Vocational Rehabilitation, 16*, 28-37.

Law, C. M., Vanderheiden, G. C. (1999). *Tests for Screening Product Designs Prior to User Testing By People with Functional Limitations*. 43rd Annual Meeting of the Human Factors and Ergonomics Society, Houston, TX.

Loricchio, D. F., and Lewis, J. R. (1991, September 2-6). *User Assessment of Standard and Reduced-Size Numeric Keypads.* Paper presented at the 35th Annual Meeting of the Human Factors and Ergonomics Society, San Francisco, CA.

Martin, D. (2000). *Doing Psychology Experiments* (pp. 366). Belmont, CA: Wadsworth/Thomson Learning.

Matlin, M. W. (1998). *Cognition*. Fort Worth, TX, Harcourt Brace College Publishers.

McNiel, J. (1997). Americans With Disabilities - Household Economic Studies (*Current Population Reports from the Survey of Income and Program Participation*). Washington, D.C.: United States Census Bureau.

Mondak, P. (2000). The Americans with Disabilities Act and Information Technology Access. *Focus on Autism and Other Developmental Disabilities* **15**(1): 43-51.

Monterey Technologies, Inc. (1996). *Resource Guide for Accessible Design of Consumer Products*. Cary, NC, EIA-EIF Committee on Product Accessibility**:** 123.

Nardi, B. (1997). The Use of Ethnographic Methods in Design and Evaluation. In M. G. Helander and T. K. Landauer and P. V. Prabhu (Eds.), *Handbook of Human-Computer Interaction* (2nd ed., pp. 361-366). Amsterdam, The Netherlands: Elsevier Science, B.V.

Newell, A., and Gregor, P. (1997). Human Computer Interfaces for People With Disabilities. In M. G. Helander and T. K. Landauer and P. V. Prabhu (Eds.), *Handbook of Human Computer Interaction* (pp. 813-824). Amsterdam, The Netherlands: Elsevier Science B.V.

Nielsen, J. (1993). *Usability Engineering*. San Diego, CA: Morgan Kaufman.

Obrien, J., Rodden, T., Rouncefield, M., Hughes, J. (1999). At Home with the Technology: An Ethnographic Study of a Set-Top-Box Trial. *ACM Transactions on Computer-Human Interaction* **6**(3): 282-308.

Pamplin College of Business Student Computer Purchase Policy Statement(2001)., [Internet / Web page]. Virginia Polytechnic Institute and State University. Available: http://www.cob.vt.edu/compute/pcrec2.htm [2001, August 2].

Pan, C. S., Shell, R.L., Schleifer, L.M. (1994). Performance Variability as an Indicator of Fatigue and Boredom in an VDT Data-Entry Task. *International Journal of Human-Computer Interaction* **6**(1): 37-45.

Philips, B. H. (1993). *Developing Interactive Guidelines for Software User Interface Design: A Case Study*. 37th Annual Meeting of the Human Factors and Ergonomics Society, Seattle, WA, HFES.

Pieper, M. (2001). Sociological Issues in HCI Design. In C. Stephanidis (Ed.), *User Interfaces For All - Concepts, Methods, and Tools* (Vol. 1, pp. 203-221). Mahwah, N.J.: Lawrence Erlbaum Associates.

Preece, J. (Ed.). (1993). *A Guide To Usability - Human Factors In Computing*. Menlo Park, CA: Addison-Wesley.

Program, A. w. D. A. T. A. (2001, July 13). *Historical Context of The Americans With Disabilities Act*, [Web page]. ADA Technical Assistance Program. Available: http://www.adata.org/whatsada-history.html [2001, September 8].

Questionnaire for User Interface Satisfaction. (2001). College Park, MD, University of Maryland.

Roscoe, S., Corl, L., Jensen, R. (1981). Flight display dynamics revisited. *Human Factors*, 23, 341-353.

Rose, A., Schneiderman, B., and Plaisant, C. (1995). *An Applied Ethnographic Method For Redesigning User Interfaces*, [Electronic conference proceeding]. Association of Computing Machinery [2001, August 28].

Rosenzweig, E. (1996). Design Guidelines for Software Products: A Common Look and Feel or a Fantasy? *interactions* **3**(5): 21-26.

Saetersdal, B. (1997). Forbidden Suffering: The Pollyanna Syndrome of the Disabled and Their Families. *Family Process* **36**(4): 431-435.

Sanders, M. S. and E. J. McCormick (1993). *Human Factors In Engineering And Design*. New York, N.Y., McGraw-Hill, Inc.

Schneiderman, B. (1998). *Designing the User Interface - Strategies for Effective Human-Computer Interaction*. Reading, MA, Addison-Wesley.

Schneiderman, B. (2000). Universal Usability: Pushing human-computer interaction research to empower every citizen. *Communications of the ACM, 43*(5), 85-91.

Serafin, C., Wen, C., Paelke, G., & Green, P. (1993). *Car Phone Usability: A Human Factors Laboratory Test.* Paper presented at the 37th Annual Meeting of the Human Factors and Ergonomics Society, Seattle, WA.

Smith, S. L., Moser, J.N. (1986). *Guidelines for Designing User Interface Software*. Bedford, MA, The MITRE Corporation.

Smith-Jackson, T., Mooney, A., and Nussbaum, M. (2001). *Survey On Technology Needs for Disabled Persons*. Unpublished Technical Report. Virginia Polytechnic Institute and State University, Blacksburg, VA.

Sommerich, C. M., Starr, H., Smith, C.A., Shivers, C. (2002). Effects of notebook computer configuration and task on user biomechanics, productivity, and comfort. *International Journal of Industrial Ergonomics* **30**: 7-31.

Stephanidis, C. (2001). User Interfaces for All: New Perspectives into Human-Computer Interaction. In C. Stephanidis (Ed.), *User Interfaces for All - Concepts, Methods, and Tools* (Vol. 1, pp. 3-17). Mahwah, NJ: Lawrence Erlbaum Associates.

Telecommunications Act of 1996, RIN 3014-AA19 (1996).

Vanderheiden, G. (1990). Thirty-Something (Million): Should They Be Exceptions? *Human Factors, 32*(4), 383-396.

Vanderheiden, G., and Henry, S. L. (2001). Everyone Interfaces. In C. Stephanidis (Ed.), *User Interfaces for All - Concepts, Methods, and Tools* (Vol. 1, pp. 115-133). Mahwah, N.J.: Lawrence Erlbaum Associates.

Virzi, R. (1997). Usability Inspection Methods. In M. G. Helander and T. K. Landauer and P. V. Prabhu (Eds.), *Handbook of Human Computer Interaction* (2nd ed., pp. 705-715). Amsterdam, The Netherlands: Elsevier Science B.V.

Volaitis, L., Chou, R., and Wiklund, M. (1998). *Consumers' Expectations of the Densities of Portable Communication Devices.* Paper presented at the 42nd Annual Meeting of the Human Factors and Ergonomics Society, Chicago, IL.

Wickens, C. D., and Hollands, J. G. (2000). *Engineering Psychology and Human Performance* (3rd ed.). Upper Saddle River, N.J.: Prentice Hall.

Williges, R. C., and Williges, B. H. (1995). Travel alternatives for mobility impaired people: The Surrogate Electronic Traveler (Set). In A. D. Edwards (Ed.), *Extra-Ordinary Human-Computer Interaction: Interfaces for Users with Disabilities* (pp. 245-261). New York: Cambridge University Press.

Appendix A - Notebook Computer Study Informed Consent Form

VIRGINIA POLYTECHNIC INSTITUTE AND STATE UNIVERSITY DEPARTMENT OF INDUSTRIAL AND SYSTEMS ENGINEERING (ISE)

Informed Consent for Participants of Investigative Projects

(note: this will be read to blind participants)
Title of Project: "Ethnographic Research on Notebook Computers"

Principal Investigators: Dr. M. A. Nussbaum, Assistant Professor, ISE
Dr. T. L. Smith-Jackson, Assistant Professor, ISE
Aaron Mooney, Graduate Research Assistant, ISE

I. THE PURPOSE OF THIS RESEARCH

The purpose of this research is to determine the usability and accessibility issues associated with the use of notebook computers among people with disabilities. Once these issues are determined, usability and accessibility recommendations for the design of notebook computers will be created.

To obtain information regarding notebook computer usability and accessibility, an ethnographic study will be conducted where participants will be asked to use two different notebook computers during a 60 day period. During this period participants will be asked to use the notebook computer at least 5 times per week. When the participant experiences difficulty, or dissatisfaction while using the machine, they will be asked to provide details on the issue. Details will be provided using a reporting process that includes a semi-automated critical incident form, as well as a description of the problem on a mini-cassette recorder (provided by the investigators). Throughout the study, investigators will contact participants to retrieve critical incident data.

II. STUDY PROCEDURES

The procedures used in this study are as follows:

Researchers will travel to a location convenient to the participant and explain the purpose and format of the research. Participants will be asked to sign a form outlining the responsibilities for the care and security of the loaned notebook computers. Informed Consent forms will also be signed at this time. Participants will be asked to use the notebook computer at least 5 times per week at a location that is convenient to them. When participants experience difficulty or dissatisfaction using the machine, participants will complete a computerized critical incident form, and provide an audio narrative. Researchers will inform participants of the need to collect critical incident data on an ongoing basis. Participants will be asked to complete a subjective evaluation form with regards to the notebook computer that was loaned to them. After 30 days of use, researchers will contact the participants to collect the notebook computers. After this first 30 day period, participants will be loaned a second notebook

computer to use for an additional 30 days. The data collection procedures will then be repeated using this second model of notebook computer.
After the second 30 day period, researchers will collect the loaned computers from the participants. Participants will be asked to complete a subjective evaluation form regarding the accessibility and usability of each computer.

III. RISKS AND BENEFITS OF THIS RESEARCH

Participation in this study will help identify usability and accessibility issues with regards to the use of notebook computers by people with disabilities.

The risks involved in this study should be minimal. There is a slight possibility of discomfort due to eyestrain or physical discomfort while using the computer. Since are being asked to use the computers as you would normally, there are no additional risks posed by this study.

IV. EXTENT OF ANONYMITY AND CONFIDENTIALITY

It is the intent of the investigators of this project to report the findings of this study. The information you provide will have your identity removed and only a participant number will identify you during analysis and any written reports of the evaluation.

V. COMPENSATION

If you decide to participate in the notebook computer portion of the study, you will be provided 2 different models of notebook computer for your personal use for a period of 60 days (30 days for each model). No other compensation will be provided, and all computer equipment provided must be returned at the completion of the study.

VI. FREEDOM TO WITHDRAW

You are free to withdraw from this study at any time for any reason without penalty. If you choose to withdraw during the study, there is no penalty. All computer equipment provided as part of this study must be returned, however, at the time of withdrawal.

VII. APPROVAL OF THIS RESEARCH

This research project has been approved, as required, by the Institutional Review Board for projects involving human participants at Virginia Polytechnic Institute and State University, and by the Department of Industrial Engineering.

VIII. PARTICIPANT RESPONSIBILITIES

I know of no reason why I cannot participate in this study. I have the responsibility of notifying the investigator at any time about a desire to discontinue participation.

IX. PARTICIPANT'S PERMISSION

Before you sign the signature page of this form, please make sure that you understand, to your complete satisfaction, the nature of the study and your rights as a participant. If you have any questions, please ask the investigator at this time. Then if you decide to participate, please sign your name above and on the following page (please repeat for your copy).

I have read a description of this study and understand the nature of the research and my rights as a participant. I hereby consent to participate, with the understanding that I may discontinue participation at any time if I choose to do so.

Signature ____________________________
Printed Name _________________________
Date ______________________________

The research team for this experiment includes Dr. M. A. Nussbaum, Assistant Professor, Dr. T. L. Smith-Jackson, Assistant Professor, and Aaron Mooney, Graduate Research Assistant. Research team members may be contacted at the following address and phone number:

Grado Department of Industrial and Systems Engineering Department
250 New Engineering Building
Virginia Tech
Blacksburg, VA 24061
(540) 231-6656

In addition, if you have detailed questions regarding your rights as a participant in University research, you may contact the following individual:

Dr. David Moore
Chair, Institutional Review Board
CVM Phase II (Pathobiology)
Virginia Tech
Blacksburg, VA 24061
(540) 231-4991

This study has been approved by the Institutional Review Board (IRB) at Virginia Polytechnic Institute and State University (IRB # 01-368).

Appendix B - Notebook Computer Study Subjective Evaluation Forms

Questionnaire for User Interface Satisfaction v. 7.0

(Used with permission from the University of Maryland)

Age: ______

Gender: ____ male

____ female

PART 1: System Experience

1.1 How long have you worked on this system?

__ less than 1 hour

__ 1 hour to less than 1 day

__ 1 day to less than 1 week

__ 1 week to less than 1 month

__ 1 month to less than 6 months

1.2 On the average, how much time do you spend per week on this system?

__ less than one hour

__ one to less than 4 hours

__ 4 to less than 10 hours

__ over 10 hours

PART 2: Past Experience

2.1 How many operating systems have you worked with?

__ none
__ 1
__ 2
__ 3-4
__ 5-6
__ more than 6

2.2 Of the following devices, software, and systems, check those that you have personally used and are familiar with:

__ computer terminal
__ personal computer
__ lap top computer
__ color monitor
__ touch screen
__ floppy drive
__ CD-ROM drive
_ keyboard
__ mouse
__ track ball
__ joy stick
__ modems
__ scanners
__ word processor
__ graphics software
__ spreadsheet software
__ database software
__ computer games
__ voice recognition
__ video editing systems
__ CAD computer aided design
__ e-mail
__ internet

PART 3: Overall User Reactions

Please mark the numbers which most appropriately reflect your impressions about using this computer system (hardware and software).
Not Applicable = NA.

3.1 Overall reactions to the system:	terrible		wonderful	
		1 2 3 4 5 6 7 8 9		NA
3.2	frustrating		satisfying	
		1 2 3 4 5 6 7 8 9		NA
3.3	dull		stimulating	
		1 2 3 4 5 6 7 8 9		NA
3.4	difficult		easy	
		1 2 3 4 5 6 7 8 9		NA
3.5	inadequate power		adequate power	
		1 2 3 4 5 6 7 8 9		NA
3.6	rigid		flexible	
		1 2 3 4 5 6 7 8 9		NA

Please write your comments about software installation here:

____

____

____

____

____

PART 4: Screen

4.1	Characters on the computer screen	hard to read	easy to read	
		1 2 3 4 5 6 7 8 9		NA
4.1.1	Image of characters	fuzzy	sharp	
		1 2 3 4 5 6 7 8 9		NA
4.1.2	Character shapes (fonts)	barely legible	very legible	
		1 2 3 4 5 6 7 8 9		NA
4.2	Highlighting on the screen	unhelpful	helpful	
		1 2 3 4 5 6 7 8 9		NA
4.2.1	Use of reverse video	unhelpful	helpful	
		1 2 3 4 5 6 7 8 9		NA
4.2.2	Use of blinking	unhelpful	helpful	
		1 2 3 4 5 6 7 8 9		NA
4.2.3	Use of bolding	unhelpful	helpful	
		1 2 3 4 5 6 7 8 9		NA
4.3	Screen layouts were helpful	never	always	
		1 2 3 4 5 6 7 8 9		NA
4.3.1	Amount of information that can be displayed on screen	inadequate	adequate	
		1 2 3 4 5 6 7 8 9		NA
4.3.2	Arrangement of information on screen	illogical	logical	
		1 2 3 4 5 6 7 8 9		NA
4.4	Sequence of screens	confusing	clear	
		1 2 3 4 5 6 7 8 9		NA
4.4.1	Next screen in a sequence	unpredictable	predictable	
		1 2 3 4 5 6 7 8 9		NA
4.4.2	Going back to the previous screen	impossible	easy	
		1 2 3 4 5 6 7 8 9		NA
4.4.3	Progression of work related tasks	confusing	clearly marked	
		1 2 3 4 5 6 7 8 9		NA
4.5	Visibility of items on screen	Poor	Excellent	
		1 2 3 4 5 6 7 8 9		NA
4.5.1	Contrast of characters on screen	Poor	Excellent	
		1 2 3 4 5 6 7 8 9		NA
4.5.2	Brightness of screen	Poor	Excellent	
		1 2 3 4 5 6 7 8 9		NA

4.5.3	Richness of colors on screen	Poor Excellent 1 2 3 4 5 6 7 8 9	NA
4.5.3	Legibility of items on screen	Poor Excellent 1 2 3 4 5 6 7 8 9	NA

Please write your comments about the screens here:

____

____

____

____

____

PART 5: Terminology and System Information

5.1	Use of terminology throughout system	inconsistent	consistent	
		1 2 3 4 5 6 7 8 9		NA
5.1.2	Work related terminology	inconsistent	consistent	
		1 2 3 4 5 6 7 8 9		NA
5.2.3	Computer terminology	inconsistent	consistent	
		1 2 3 4 5 6 7 8 9		NA
5.2	Terminology relates well to the work you are doing?	always	never	
		1 2 3 4 5 6 7 8 9		NA
5.2.1	Computer terminology is used	too frequently	appropriately	
		1 2 3 4 5 6 7 8 9		NA
5.2.2	Terminology on the screen	ambiguous	precise	
		1 2 3 4 5 6 7 8 9		NA
5.3	Messages which appear on screen	inconsistent	consistent	
		1 2 3 4 5 6 7 8 9		NA
5.3.1	Position of instructions on the screen	inconsistent	Consistent	
		1 2 3 4 5 6 7 8 9		NA
5.4	Messages which appear on screen	confusing	clear	
		1 2 3 4 5 6 7 8 9		NA
5.4.1	Instructions for commands or functions	confusing	clear	
		1 2 3 4 5 6 7 8 9		NA
5.4.2	Instructions for correcting errors	confusing	clear	
		1 2 3 4 5 6 7 8 9		NA
5.5	Computer keeps you informed about what it is doing	never	always	
		1 2 3 4 5 6 7 8 9		NA
5.5.1	Animated cursors keep you informed	never	always	
		1 2 3 4 5 6 7 8 9		NA
5.5. 2	Performing an operation leads to a predictable result	never	always	
		1 2 3 4 5 6 7 8 9		NA
5.5.3	Controlling amount of feedback	impossible	easy	
		1 2 3 4 5 6 7 8 9		NA
5.5.4	Length of delay between operation	unacceptable	acceptable	
		1 2 3 4 5 6 7 8 9		NA

5.6 Error messages	unhelpful — helpful	1 2 3 4 5 6 7 8 9	NA
5.6.1 Error messages clarify the problem	never — always	1 2 3 4 5 6 7 8 9	NA
5.6.2 Phrasing of error messages	unpleasant — pleasant	1 2 3 4 5 6 7 8 9	NA

Please write your comments about terminology and system information here:

____

____

____

____

____

PART 6: Learning

6.1	Learning to operate the system	difficult 1 2 3 4 5 6 7 8 9 easy		NA
6.1.1	Getting started	difficult 1 2 3 4 5 6 7 8 9 easy		NA
6.1.2	Learning advanced features	difficult 1 2 3 4 5 6 7 8 9 easy		NA
6.1.3	Time to learn to use the system	slow 1 2 3 4 5 6 7 8 9 fast		NA
6.2	Exploration of features by trial and error	discouraging 1 2 3 4 5 6 7 8 9 encouraging		NA
6.2.1	Exploration of features	risky 1 2 3 4 5 6 7 8 9 safe		NA
6.2.2	Discovering new features	difficult 1 2 3 4 5 6 7 8 9 easy		NA
6.3	Remembering names and use of commands	difficult 1 2 3 4 5 6 7 8 9 easy		NA
6.3.1	Remembering specific rules about entering commands	difficult 1 2 3 4 5h 6 7 8 9 easy		NA
6.4	Tasks can be performed in a straight-forward manner	never 1 2 3 4 5 6 7 8 9 always		NA
6.4.1	Number of steps per task	too many 1 2 3 4 5 6 7 8 9 just right		NA
6.4.2	Steps to complete a task follow a logical sequence	never 1 2 3 4 5 6 7 8 9 always		NA
6.4.3	Feedback on the completion of of steps	clear 1 2 3 4 5 6 7 8 9 unclear		NA

Please write your comments about learning here:

____

____

____

____

____

PART 7: System Capabilities

7.1	System speed	too slow ... fast enough 1 2 3 4 5 6 7 8 9	NA
7.1.1	Response time for most operations	too slow ... fast enough 1 2 3 4 5 6 7 8 9	NA
7.1.2	Rate information is displayed	too slow ... fast enough 1 2 3 4 5 6 7 8 9	NA
7.2	The system is reliable	never ... always 1 2 3 4 5 6 7 8 9	NA
7.2.1	Operations are	undependable ... dependable 1 2 3 4 5 6 7 8 9	NA
7.2.2	System failures occur	frequently ... seldom 1 2 3 4 5 6 7 8 9	NA
7.2.3	System warns you about potential problems	never ... always 1 2 3 4 5 6 7 8 9	NA
7.3	System tends to be	quiet ... noisy 1 2 3 4 5 6 7 8 9	NA
7.3.1	Mechanical devices such as fans, disks, and printers	noisy ... quiet 1 2 3 4 5 6 7 8 9	NA
7.3.2	Computer generated sounds are	annoying ... pleasant 1 2 3 4 5 6 7 8 9	NA
7.4	Correcting your mistakes	difficult ... easy 1 2 3 4 5 6 7 8 9	NA
7.4.1	Correcting typos	complex ... simple 1 2 3 4 5 6 7 8 9	NA
7.4.2	Ability to undo operations	inadequate ... adequate 1 2 3 4 5 6 7 8 9	NA
7.5	Ease of operation depends on your level of experience	never ... always 1 2 3 4 5 6 7 8 9	NA
7.5.1	You can accomplish tasks knowing only a few commands	with difficulty ... easily 1 2 3 4 5 6 7 8 9	NA
7.5.2	You can use features/shortcuts	with difficulty ... easily 1 2 3 4 5 6 7 8 9	NA
7.6	The computer interferes with the use of accessibility software	never ... always 1 2 3 4 5 6 7 8 9	NA

7.6.1 Web browsing with accessibility Software is — difficult 1 2 3 4 5 6 7 8 9 easy — NA

7.6.2 Using email with accessibility Software is — difficult 1 2 3 4 5 6 7 8 9 easy — NA

7.6.3 Using word processing applications with accessibility software is — difficult 1 2 3 4 5 6 7 8 9 easy — NA

Please write your comments about system capabilities here:

____

____

____

____

PART 8: Technical Manuals and On-line help

8.1	Technical manuals are	confusing	clear	
		1 2 3 4 5 6 7 8 9		NA
8.1.1	The terminology used in the manual	confusing	clear	
		1 2 3 4 5 6 7 8 9		NA
8.2	Information from the manual is easily understood	never	always	
		1 2 3 4 5 6 7 8 9		NA
8.2.1	Finding a solution to a problem using the manual	impossible	easy	
		1 2 3 4 5 6 7 8 9		NA
8.3	Amount of help given	inadequate	adequate	
		1 2 3 4 5 6 7 8 9		NA
8.3.1	Placement of help messages on the screen	confusing	clear	
		1 2 3 4 5 6 7 8 9		NA
8.3.2	Accessing help messages	difficult	easy	
		1 2 3 4 5 6 7 8 9		NA
8.3.3	Content of on-line help messages	confusing	clear	
		1 2 3 4 5 6 7 8 9		NA
8.3.4	Amount of help given	inadequate	adequate	
		1 2 3 4 5 6 7 8 9		NA
8.3.5	Help defines specific aspects of the system	inadequately	adequately	
		1 2 3 4 5 6 7 8 9		NA
8.3.6	Finding specific information using the on-line help	difficult	easy	
		1 2 3 4 5 6 7 8 9		NA
8.3.7	On-line help	useless	helpful	
		1 2 3 4 5 6 7 8 9		NA

Please write your comments about technical manuals and on-line help here:

____

____

____

____

____

PART 9: On-line Tutorials

9.1 Tutorial was	useless	helpful	
	1 2 3 4 5 6 7 8 9		NA
9.1.1 Accessing on-line tutorial	difficult	easy	
	1 2 3 4 5 6 7 8 9		NA
9.2 Maneuvering through the tutorial was	difficult	easy	
	1 2 3 4 5 6 7 8 9		NA
9.2.1 Tutorial is meaningfully structured	never	always	
	1 2 3 4 5 6 7 8 9		NA
9.2.2 The speed of presentation was	unacceptable	acceptable	
	1 2 3 4 5 6 7 8 9		NA
9.3 Tutorial content was	useless	helpful	
	1 2 3 4 5 6 7 8 9		NA
9.3.1 Information for specific aspects of the system were complete and informative	never	always	
	1 2 3 4 5 6 7 8 9		NA
9.3.2 Information was concise and to the point	never	always	
	1 2 3 4 5 6 7 8 9		NA
9.4 Tasks can be completed	with difficulty	easily	
	1 2 3 4 5 6 7 8 9		NA
9.4.1 Instructions given for completing tasks	confusing	clear	
	1 2 3 4 5 6 7 8 9		NA
9.4.2 Time given to perform tasks	inadequate	adequate	
	1 2 3 4 5 6 7 8 9		NA
9.5 Learning to operate the system using the tutorial was	difficult	easy	
	1 2 3 4 5 6 7 8 9		NA
9.5.1 Completing system tasks after using only the tutorial	difficult	easy	
	1 2 3 4 5 6 7 8 9		NA

Please write your comments about on-line tutorials here:

____

____

____

____

____

PART 10: Multimedia

10.1 Quality of still pictures/photographs	bad	good	
	1 2 3 4 5 6 7 8 9		NA
10.1.1 Pictures/Photos	fuzzy	clear	
	1 2 3 4 5 6 7 8 9		NA
10.1.2 Picture/Photo brightness	dim	bright	
	1 2 3 4 5 6 7 8 9		NA
10.2 Quality of movies	bad	good	
	1 2 3 4 5 6 7 8 9		NA
10.2.1 Focus of movie images	fuzzy	clear	
	1 2 3 4 5 6 7 8 9		NA
10.2.2 Brightness of movie images	dim	bright	
	1 2 3 4 5 6 7 8 9		NA
10.2.3 Movie window size is adequate	never	always	
	1 2 3 4 5 6 7 8 9		NA
10.3 Sound output	inaudible	audible	
	1 2 3 4 5 6 7 8 9		NA
10.3.1 Sound output	choppy	smooth	
	1 2 3 4 5 6 7 8 9		NA
10.3.2 Sound output	garbled	clear	
	1 2 3 4 5 6 7 8 9		NA
10.4 Colors used are	unnatural	natural	
	1 2 3 4 5 6 7 8 9		NA
10.4.1 Amount of colors available	inadequate	adequate	
	1 2 3 4 5 6 7 8 9		NA

Please write your comments about multimedia here:

____

____

____

____

____

PART 11: Teleconferencing

11.1	Setting up for conference	difficult	easy	
		1 2 3 4 5 6 7 8 9		NA
11.1.1	Time for establishing the connections to others	too long	just right	
		1 2 3 4 5 6 7 8 9		NA
11.1.2	Number of connections possible	too few	enough	
		1 2 3 4 5 6 7 8 9		NA
11.2	Arrangement of windows showing connecting groups	confusing	clear	
		1 2 3 4 5 6 7 8 9		NA
11.2.1	Window with view of your own group is of appropriate size	never	always	
		1 2 3 4 5 6 7 8 9		NA
11.2.2	Window(s) with view of connecting group(s) is of appropriate size	never	always	
		1 2 3 4 5 6 7 8 9		NA
11.3	Determining the focus of attention during conference was	confusing	clear	
		1 2 3 4 5 6 7 8 9		NA
11.3.1	Telling who is speaking	difficult	easy	
		1 2 3 4 5 6 7 8 9		NA
11.4	Video image flow	choppy	smooth	
		1 2 3 4 5 6 7 8 9		NA
11.4.1	Focus of video image	fuzzy	clear	
		1 2 3 4 5 6 7 8 9		NA
11.5	Audio output	audible	inaudible	
		1 2 3 4 5 6 7 8 9		NA
11.5.1	Audio is in sync with video images	never	always	
		1 2 3 4 5 6 7 8 9		NA
11.6	Exchanging data	difficult	easy	
		1 2 3 4 5 6 7 8 9		NA
11.6.1	Transmitting files	difficult	easy	
		1 2 3 4 5 6 7 8 9		NA
11.6.2	Retrieving files	difficult	easy	
		1 2 3 4 5 6 7 8 9		NA
11.6.3	Using on-line chat	difficult	easy	
		1 2 3 4 5 6 7 8 9		NA
11.6.4	Using shared workspace	difficult	easy	
		1 2 3 4 5 6 7 8 9		NA

Please write your comments about teleconferencing here:

____

____

____

____

____

PART 12: Software Installation

12.1 Speed of installation	slow		fast	
	1 2 3 4 5 6 7 8 9			NA
12.2 Customization	difficult		easy	
	1 2 3 4 5 6 7 8 9			NA
12.2.1 Installing only the software you want	confusing		clear	
	1 2 3 4 5 6 7 8 9			NA
12.2.2 Removing old software versions	with difficulty		automatic	
	1 2 3 4 5 6 7 8 9			NA
12.3 Informs you of its progress	never		always	
	1 2 3 4 5 6 7 8 9			NA
12.4 Gives a meaningful explanation when failures occur	never		always	
	1 2 3 4 5 6 7 8 9			NA
12.5 Installation of accessibility software	difficult		easy	
	1 2 3 4 5 6 7 8 9			NA

Please write your comments about software installation here:

____

____

____

____

____

Part 13: Keyboard Use

13.1 Use of keyboard	Difficult	Easy	
	1 2 3 4 5 6 7 8 9		NA
13.2 Key activation force	difficult	easy	
	1 2 3 4 5 6 7 8 9		NA
13.3 Key size	Too small	Too large	
	1 2 3 4 5 6 7 8 9		NA
13.4 Distance between keys	Too small	Too large	
	1 2 3 4 5 6 7 8 9		NA
13.5 Height of keys	Too low	Too high	
	1 2 3 4 5 6 7 8 9		NA
13.6 Key shape	Poor	Excellent	
	1 2 3 4 5 6 7 8 9		NA
13.7 Key color	Poor	Excellent	
	1 2 3 4 5 6 7 8 9		NA
13.8 Primary character / key contrast	Poor	Excellent	
	1 2 3 4 5 6 7 8 9		NA
13.9 Icon / key contrast	Poor	Excellent	
	1 2 3 4 5 6 7 8 9		NA
13.10 Key activation sound	Poor	Excellent	
	1 2 3 4 5 6 7 8 9		NA
13.11 Tactile feel of keys	Poor	Excellent	
	1 2 3 4 5 6 7 8 9		NA

Please write your comments about keyboard use here:

____

____

____

____

____

Part 14: Pointing Device Use

14.1	Use of pointing device	Difficult Easy 1 2 3 4 5 6 7 8 9	NA
14.2	Pointing device activation force	difficult easy 1 2 3 4 5 6 7 8 9	NA
14.3	Pointing device size	Too small Too large 1 2 3 4 5 6 7 8 9	NA
14.4	Pointing device texture	Poor excellent 1 2 3 4 5 6 7 8 9	NA
14.5	Height of pointing device	Too low Too high 1 2 3 4 5 6 7 8 9	NA
14.6	Shape of pointing device	Poor excellent 1 2 3 4 5 6 7 8 9	NA
14.7	Pointing device color	Poor Excellent 1 2 3 4 5 6 7 8 9	NA
14.8	Pointing device color / key contrast	Poor Excellent 1 2 3 4 5 6 7 8 9	NA
14.9	Ability to control cursor with pointing device	Poor Excellent 1 2 3 4 5 6 7 8 9	NA
14.10	Size of pointing device buttons	Poor Excellent 1 2 3 4 5 6 7 8 9	NA
14.11	Activation force of pointing device buttons	Poor Excellent 1 2 3 4 5 6 7 8 9	NA

Please write your comments about pointing device use here:

____

____

____

____

____

Part 15: Miscellaneous Hardware

15.1	Ability to open and close screen	Difficult Easy 1 2 3 4 5 6 7 8 9	NA
15.2	Location of power key	Poor Excellent 1 2 3 4 5 6 7 8 9	NA
15.3	Ease of finding power key	Difficult Easy 1 2 3 4 5 6 7 8 9	NA
15.4	Location of power supply connection	Poor Excellent 1 2 3 4 5 6 7 8 9	NA
15.5	Ease of finding power supply connection	Difficult Easy 1 2 3 4 5 6 7 8 9	NA
15.6	Location of USB port	Poor excellent 1 2 3 4 5 6 7 8 9	NA
15.7	Ease of plugging in USB devices	Difficult Easy 1 2 3 4 5 6 7 8 9	NA
15.8	Location of telephone / Ethernet connections	Poor Excellent 1 2 3 4 5 6 7 8 9	NA
15.9	Ease of using telephone / Ethernet connection	Difficult Easy 1 2 3 4 5 6 7 8 9	NA
15.10	Location of external pointing device port	Poor Excellent 1 2 3 4 5 6 7 8 9	NA
15.11	Ease of using external pointing device port	Difficult Easy 1 2 3 4 5 6 7 8 9	NA
15.12	Location of microphone / speaker ports	Poor Excellent 1 2 3 4 5 6 7 8 9	NA
15.13	Ease of use of microphone / speaker ports	Difficult Easy 1 2 3 4 5 6 7 8 9	NA
15.14	Location of CD drive	Poor Excellent 1 2 3 4 5 6 7 8 9	NA
15.15	Ease of use of CD drive	Difficult Easy 1 2 3 4 5 6 7 8 9	NA
15.16	Location of Smart Media card slots	Poor Excellent 1 2 3 4 5 6 7 8 9	NA
15.17	Ease of inserting Smart Media card	Difficult Easy 1 2 3 4 5 6 7 8 9	NA

15.18	Ease of removing Smart Media card	Difficult Easy 1 2 3 4 5 6 7 8 9	NA
15.19	Ease of plugging in power supply to Computer.	Difficult Easy 1 2 3 4 5 6 7 8 9	NA
15.20	Location of speaker volume control	Poor Excellent 1 2 3 4 5 6 7 8 9	NA
15.21	Ease of using speaker volume control	Difficult Easy 1 2 3 4 5 6 7 8 9	NA

Please write your comments about miscellaneous hardware here:

____

____

____

____

____

Part 16: Accessibility

16.1	Latch(es) to open screen	Not accessible 1 2 3 4 5 6 7 8 9	Very accessible	NA
16.2	Power key	Not accessible 1 2 3 4 5 6 7 8 9	Very accessible	NA
16.3	Location of power supply connection	Not accessible 1 2 3 4 5 6 7 8 9	Very accessible	NA
16.4	Location of USB port	Not accessible 1 2 3 4 5 6 7 8 9	Very accessible	NA
16.5	Use of USB port	Not accessible 1 2 3 4 5 6 7 8 9	Very accessible	NA
16.6	Location of telephone / Ethernet connections	Not accessible 1 2 3 4 5 6 7 8 9	Very accessible	NA
16.7	Use of Ethernet / telephone connections	Not accessible 1 2 3 4 5 6 7 8 9	Very accessible	NA
16.8	Location of external pointing device port	Not accessible 1 2 3 4 5 6 7 8 9	Very accessible	NA
16.9	Use of external pointing device port	Not accessible 1 2 3 4 5 6 7 8 9	Very accessible	NA
16.10	Location of microphone / speaker ports	Not accessible 1 2 3 4 5 6 7 8 9	Very accessible	NA
16.11	Use of microphone / speaker ports	Not accessible 1 2 3 4 5 6 7 8 9	Very accessible	NA
16.12	Location of CD drive	Not accessible 1 2 3 4 5 6 7 8 9	Very accessible	NA
16.13	Use of CD drive	Not accessible 1 2 3 4 5 6 7 8 9	Very accessible	NA
16.14	Location of Smart media card slots	Not accessible 1 2 3 4 5 6 7 8 9	Very accessible	NA
16.15	Insertion of Smart Media card	Not accessible 1 2 3 4 5 6 7 8 9	Very accessible	NA
16.16	Removal of Smart Media card	Not accessible 1 2 3 4 5 6 7 8 9	Very accessible	NA
16.17	Use of brightness adjustment	Not accessible 1 2 3 4 5 6 7 8 9	Very accessible	NA
16.18	Use of speaker volume control	Not accessible	Very accessible	

1 2 3 4 5 6 7 8 9 NA

Please write your comments about accessibility here:

____

____

____

____

____

Appendix C - Cellular Telephone Study Informed Consent Form

VIRGINIA POLYTECHNIC INSTITUTE AND STATE UNIVERSITY DEPARTMENT OF INDUSTRIAL AND SYSTEMS ENGINEERING (ISE)

Informed Consent for Participants of Investigative Projects

(note: this will be read to blind participants)
Title of Project: "Performance and Usability Evaluation of Cellular Phones for Persons with Disabilities"

Principal Investigators: Dr. M. A. Nussbaum, Assistant Professor, ISE
Dr. T. L. Smith-Jackson, Assistant Professor, ISE
Aaron Mooney, Graduate Research Assistant, ISE

I. THE PURPOSE OF THIS RESEARCH

The purpose of this research is to determine which keypad and display attributes of cellular phones provide the best performance and subjective evaluations for disabled users.

Four specific cellular phone attributes will be studied experimentally: key design; location of power-up and send/end key; confirmation of key operations; and, visibility (readability) of letters on an LCD. To obtain this information, participants will be asked to perform common phone-related tasks on phones provided by the Toshiba Corporation of Japan. The performance of each participant will be videotaped (including audio) for further analysis. Additionally, participants will be asked to complete subjective evaluations of the phones to determine usability, accessibility, etc.

The results of this study will be used to develop recommendations for the usability and accessibility of cellular phones.

II. STUDY PROCEDURES

The procedures used in this study are as follows:

Cellular phone performance testing will be conducted at Virginia Tech. If participants require taxi service to campus, they will be compensated for this expense. Participants will be given a brief introduction to the purpose and format of the testing, and the Informed Consent form will be read and signed. The participant will be given time to familiarize themselves with the cellular phone models available for testing. Participants will then be asked to perform several common phone-related tasks (such as dialing, etc.). The entire test will be videotaped, including audio. In order to maintain anonymity, the participants' face will not be included in the video. After participants complete performance testing, they will be asked to complete subjective evaluations on the test attributes, as well as on the prototype usability in general. Once all evaluations are complete, participants will be paid and transportation home will be arranged.

III. RISKS AND BENEFITS OF THIS RESEARCH

Participation in this study will help identify what cellular phone attributes lead to increased usability and accessibility for users with disabilities. There is minimal risk involved in this study.

The benefits of this study will be the creation of usability and accessibility recommendations for the use of cellular phones by people with disabilities. It is hoped that these recommendations will be used in the future design and manufacture of cellular phones.

IV. EXTENT OF ANONYMITY AND CONFIDENTIALITY

It is the intent of the investigators of this project to report the findings of this study. The information you provide will have your name removed and only a subject number will identify you during analysis and any written reports of the evaluation.

V. COMPENSATION

If you decide to participate in this study, you will be paid $12.00 per hour for the time you participate. Each survey is expected to last approximately 3 hours total. You will be paid at the conclusion of the survey session.

VI. FREEDOM TO WITHDRAW

You are free to withdraw from this study at any time for any reason without penalty. If you choose to withdraw during the study, you will be compensated for the portion of the testing which has been completed.

VII. APPROVAL OF THIS RESEARCH

This research project has been approved, as required, by the Institutional Review Board for projects involving human participants at Virginia Polytechnic Institute and State University, and by the Department of Industrial Engineering.

VIII. PARTICIPANT RESPONSIBILITIES

I know of no reason why I cannot participate in this study. I agree to notify the investigator at any time about a desire to discontinue participation.

IX. PARTICIPANT'S PERMISSION

Before you sign the signature page of this form, please make sure that you understand, to your complete satisfaction, the nature of the study and your rights as a participant. If you have any questions, please ask the investigator at this time. If you then decide to participate, please sign your name above and on the following page (please repeat for your copy).

I have read a description of this study and understand the nature of the research and my rights as a participant. I hereby consent to participate, with the understanding that I may discontinue participation at any time if I choose to do so.

Signature ___________________________

Printed Name ________________________

Date ____________________________

The research team for this experiment includes Dr. M. A. Nussbaum, Assistant Professor, Dr. T. L. Smith-Jackson, Assistant Professor, and Aaron Mooney, Graduate Research Assistant.. Research team members may be contacted at the following address and phone number:

Grado Department of Industrial and Systems Engineering Department
250 New Engineering Building
Virginia Tech
Blacksburg, VA 24061
(540) 231-6656

In addition, if you have detailed questions regarding your rights as a participant in University research, you may contact the following individual:

Dr. David Moore
Chair, Institutional Review Board
CVM Phase II (Pathobiology)
Virginia Tech
Blacksburg, VA 24061
(540) 231-4991

This study has been approved by the Institutional Review Board (IRB) at Virginia Polytechnic Institute and State University (IRB # 01-369).

Appendix D - Cellular Telephone Study Subjective Evaluation Forms

Cellular Telephone Subjective Evaluation Form

Participant #: _____ Age: _____ Gender: _____

Disability (please circle) :

| None | Total blindness | Legal blindness |

| Severe Upper Extremity Physical | Minor Upper Extremity Physical |

Instructions:

Please rank each phone in order of preference based solely on the attribute in question. Mark the number 1 above the one you preferred the most, the number 2 above your second choice, and so forth.

1. Please rank the telephones listed below in the order of preference, based solely on lateral key distance:

____ ____ ____ ____

A1 A2 A3 A4

2. Why did you prefer the telephone you ranked the highest in the previous question?

3. Please rank the telephones listed below in the order of preference, based solely on keystroke:

____ ____ ____

B1 B2 B3

4. Please rank the telephones listed below in the order of preference, based solely on key protrusion:

____	____	____
C1	C2	C3

5. Please rank the telephones listed below in the order of preference, based solely on position of power key:

____	____	____	____	____	____	____	____
X1	X2	X3	X4	X5	X6	X7	X8

6. Why did you prefer the telephone you ranked the highest in the previous question?

7. Please rank the telephones listed below in the order of preference, based solely on position of the “Send” and “End” keys:

____	____	____	____	____	____	____	____
X1	X2	X3	X4	X5	X6	X7	X8

8. Why did you prefer the telephone you ranked the highest in the previous question?

9. Please rank the telephones listed below in the order of preference, based solely on font type:

____	____	____	____	____	____
Z1	Z2	Z3	Z4	Z5	Z6

10. Please rank your preference for the size of the alphabetic characters:

______	_______	______
Small	Medium	Large

Lateral Pitch

---

1. Please rate the ease of use of the lateral pitch on telephone A1:

|________|________|________|________|________|________|
1 2 3 4 5 6 7
Difficult Easy

2. Please rate the accessibility of lateral pitch on telephone A1:

|________|________|________|________|________|________|
1 2 3 4 5 6 7
Not Accessible Very Accessible

3. Please rate the ease of use of the lateral pitch on telephone A2:

|________|________|________|________|________|________|
1 2 3 4 5 6 7
Difficult Easy

4. Please rate the accessibility of lateral pitch on telephone A2:

|________|________|________|________|________|________|
1 2 3 4 5 6 7
Not Accessible Very Accessible

5. Please rate the ease of use of the lateral pitch on telephone A3:

|________|________|________|________|________|________|
1 2 3 4 5 6 7
Difficult Easy

6. Please rate the accessibility of lateral pitch on telephone A3:

|________|________|________|________|________|________|
1 2 3 4 5 6 7
Not Accessible Very Accessible

7. Please rate the ease of use of the lateral pitch on telephone A4:

|________|________|________|________|________|________|
1 2 3 4 5 6 7
Difficult Easy

8. Please rate the accessibility of lateral pitch on telephone A4:

|________|________|________|________|________|________|
1 2 3 4 5 6 7
Not Accessible Very Accessible

Keystroke:

1. Please rate the ease of use of the keystroke on telephone B1:

|________|________|________|________|________|________|
1 2 3 4 5 6 7
Difficult Easy

2. Please rate the accessibility of keystroke on telephone B1:

|________|________|________|________|________|________|
1 2 3 4 5 6 7
Not Accessible Very Accessible

3. Please rate the ease of use of the keystroke on telephone B2:

|________|________|________|________|________|________|
1 2 3 4 5 6 7
Difficult Easy

4. Please rate the accessibility of keystroke on telephone B3:

|________|________|________|________|________|________|
1 2 3 4 5 6 7
Not Accessible Very Accessible

Please rate the ease of use of the keystroke on telephone B3:

|________|________|________|________|________|________|
1 2 3 4 5 6 7
Difficult Easy

6. Please rate the accessibility of keystroke on telephone B3:

|________|________|________|________|________|________|
1 2 3 4 5 6 7
Not Accessible Very Accessible

Key Protrusion:

1. Please rate the ease of use of the key protrusion on telephone C1:

|________|________|________|________|________|________|
1 2 3 4 5 6 7
Difficult Easy

2. Please rate the accessibility of key protrusion on telephone C1:

|________|________|________|________|________|________|
1 2 3 4 5 6 7
Not Accessible Very Accessible

3. Please rate the ease of use of the key protrusion on telephone C2:

|________|________|________|________|________|________|
1 2 3 4 5 6 7
Difficult Easy

4. Please rate the accessibility of key protrusion on telephone C2:

|________|________|________|________|________|________|
1 2 3 4 5 6 7
Not Accessible Very Accessible

5. Please rate the ease of use of the key protrusion on telephone C3:

|________|________|________|________|________|________|
1 2 3 4 5 6 7
Difficult Easy

6. Please rate the accessibility of key protrusion on telephone C3:

|________|________|________|________|________|________|
1 2 3 4 5 6 7
Not Accessible Very Accessible

Power Key:

1. Please rate the ease of use of the power key on telephone X1:

|________|________|________|________|________|________|________|
1 2 3 4 5 6 7
Difficult Easy

2. Please rate the accessibility of power key on telephone X1:

|________|________|________|________|________|________|________|
1 2 3 4 5 6 7
Not Accessible Very Accessible

3. Please rate the ease of use of the power key on telephone X2:

|________|________|________|________|________|________|________|
1 2 3 4 5 6 7
Difficult Easy

4. Please rate the accessibility of power key on telephone X2:

|________|________|________|________|________|________|________|
1 2 3 4 5 6 7
Not Accessible Very Accessible

5. Please rate the ease of use of the power key on telephone X3

|________|________|________|________|________|________|
1 2 3 4 5 6 7
Difficult Easy

6. Please rate the accessibility of power key on telephone X3:

|________|________|________|________|________|________|
1 2 3 4 5 6 7
Not Accessible Very Accessible

7. Please rate the ease of use of the power key on telephone X4:

|________|________|________|________|________|________|
1 2 3 4 5 6 7
Difficult Easy

8. Please rate the accessibility of power key on telephone X4:

|________|________|________|________|________|________|
1 2 3 4 5 6 7
Not Accessible Very Accessible

9. Please rate the ease of use of the power key on telephone X5:

|________|________|________|________|________|________|
1 2 3 4 5 6 7
Difficult Easy

10. Please rate the accessibility of power key on telephone X5:

|________|________|________|________|________|________|
1 2 3 4 5 6 7
Not Accessible Very Accessible

11. Please rate the ease of use of the power key on telephone X6:

|________|________|________|________|________|________|
1 2 3 4 5 6 7
Difficult Easy

12. Please rate the accessibility of power key on telephone X6:

|________|________|________|________|________|________|
1 2 3 4 5 6 7
Not Accessible Very Accessible

13. Please rate the ease of use of the power key on telephone X7:

|________|________|________|________|________|________|

1 2 3 4 5 6 7

Difficult Easy

14. Please rate the accessibility of power key on telephone X7:

|________|________|________|________|________|________|

1 2 3 4 5 6 7

Not Accessible Very Accessible

15. Please rate the ease of use of the power key on telephone X8:

|________|________|________|________|________|________|

1 2 3 4 5 6 7

Difficult Easy

16. Please rate the accessibility of power key on telephone X8:

|________|________|________|________|________|________|

1 2 3 4 5 6 7

Not Accessible Very Accessible

Send and End keys:

1. Please rate the ease of use of the “Send” and “End” keys on telephone X1:

|________|________|________|________|________|________|
1 2 3 4 5 6 7
Difficult Easy

2. Please rate the accessibility of the “Send” and “End” keys on telephone X1:

|________|________|________|________|________|________|
1 2 3 4 5 6 7
Not Accessible Very
Accessible

3. Please rate the ease of use of the “Send” and “End” keys on telephone X2:

|________|________|________|________|________|________|
1 2 3 4 5 6 7
Difficult Easy

4. Please rate the accessibility of the “Send” and “End” keys on telephone X2:

|________|________|________|________|________|________|
1 2 3 4 5 6 7
Not Accessible Very Accessible

5. Please rate the ease of use of the “Send” and “End” keys on telephone X3:

|________|________|________|________|________|________|
1 2 3 4 5 6 7
Difficult Easy

6. Please rate the accessibility of the “Send” and “End” keys on telephone X3:

|________|________|________|________|________|________|
1 2 3 4 5 6 7
Not Accessible Very Accessible

7. Please rate the ease of use of the “Send” and “End” keys on telephone X4:

|________|________|________|________|________|________|
1 2 3 4 5 6 7
Difficult Easy

8. Please rate the accessibility of the “Send” and “End” keys on telephone X4:

|________|________|________|________|________|________|
1 2 3 4 5 6 7
Not Accessible Very Accessible

9. Please rate the ease of use of the “Send” and “End” keys on telephone X5:

|________|________|________|________|________|________|
1 2 3 4 5 6 7
Difficult Easy

10. Please rate the accessibility of the “Send” and “End” keys on telephone X5:

|________|________|________|________|________|________|
1 2 3 4 5 6 7
Not Accessible Very Accessible

11. Please rate the ease of use of the “Send” and “End” keys on telephone X6:

|________|________|________|________|________|________|
1 2 3 4 5 6 7
Difficult Easy

12. Please rate the accessibility of the “Send” and “End” keys on telephone X6:

|________|________|________|________|________|________|
1 2 3 4 5 6 7
Not Accessible Very Accessible

13. Please rate the ease of use of the “Send” and “End” keys on telephone X7:

|________|________|________|________|________|________|

1 2 3 4 5 6 7

Difficult Easy

14. Please rate the accessibility of the “Send” and “End” keys on telephone X7:

|________|________|________|________|________|________|

1 2 3 4 5 6 7

Not Accessible Very Accessible

14. Please rate the ease of use of the “Send” and “End” keys on telephone X8:

|________|________|________|________|________|________|

1 2 3 4 5 6 7

Difficult Easy

15. Please rate the accessibility of the “Send” and “End” keys on telephone X8:

|________|________|________|________|________|________|

1 2 3 4 5 6 7

Not Accessible Very Accessible

Numeric Display Font:

1. Given the fact that a larger display font will limit the amount of information presented on the display, which is more important to you?

    __ Font size

    __ Amount of information presented

2. Please rate the ease of use of the display font on telephone Z1:

|________|________|________|________|________|________|
1 2 3 4 5 6 7
Difficult Easy

3. Please rate the accessibility of the display font on telephone Z1:

|________|________|________|________|________|________|
1 2 3 4 5 6 7
Not Accessible Very Accessible

4. Please rate the ease of use of the display font on telephone Z2:

|________|________|________|________|________|________|
1 2 3 4 5 6 7
Difficult Easy

5. Please rate the accessibility of the display font on telephone Z2:

|________|________|________|________|________|________|
1 2 3 4 5 6 7
Not Accessible Very Accessible

5. Please rate the ease of use of the display font on telephone Z3:

|________|________|________|________|________|________|
1 2 3 4 5 6 7
Difficult Easy

6. Please rate the accessibility of the display font on telephone Z3:

|________|________|________|________|________|________|
1 2 3 4 5 6 7
Not Accessible Very Accessible

7. Please rate the ease of use of the display font on telephone Z4:

|________|________|________|________|________|________|
1 2 3 4 5 6 7
Difficult Easy

8. Please rate the accessibility of the display font on telephone Z4:

|________|________|________|________|________|________|
1 2 3 4 5 6 7
Not Accessible Very Accessible

9. Please rate the ease of use of the display font on telephone Z5:

|________|________|________|________|________|________|
1 2 3 4 5 6 7
Difficult Easy

10. Please rate the accessibility of the display font on telephone Z5:

|________|________|________|________|________|________|
1 2 3 4 5 6 7
Not Accessible Very Accessible

11. Please rate the ease of use of the display font on telephone Z6:

|________|________|________|________|________|________|
1 2 3 4 5 6 7
Difficult Easy

12. Please rate the accessibility of the display font on telephone Z6:

|________|________|________|________|________|________|
1 2 3 4 5 6 7
Not Accessible Very Accessible

Alphabetic Display Font:

1. Please rate the ease of use of the small size display font:

|________|________|________|________|________|________|

1 2 3 4 5 6 7

Difficult Easy

2. Please rate the accessibility of the small size display font:

|________|________|________|________|________|________|

1 2 3 4 5 6 7

Not Accessible Very Accessible

3. Please rate the ease of use of the medium size display font:

|________|________|________|________|________|________|

1 2 3 4 5 6 7

Difficult Easy

4. Please rate the accessibility of the medium size display font:

|________|________|________|________|________|________|

1 2 3 4 5 6 7

Not Accessible Very Accessible

5. Please rate the ease of use of the large size display font:

|________|________|________|________|________|________|

1	2	3	4	5	6	7
Difficult						Easy

6. Please rate the accessibility of the large size display font:

|________|________|________|________|________|________|

1	2	3	4	5	6	7
Not Accessible						Very Accessible

Vita

Aaron Mooney received a Bachelor of Science degree in Environmental Health from Colorado State University in 1995. Upon graduation he was employed by the Hewlett-Packard Company as an Environmental Health and Safety Coordinator. In this role he performed ergonomic evaluations, safety audits, emergency response risk assessments, provided training, and implemented process improvements whenever necessary.

Following his work at Hewlett-Packard, he returned to Virginia Tech to further develop his interests in Human Factors and Ergonomics Engineering. While at Virginia Tech he worked as a teaching assistant for an industrial ergonomics course, and then worked with people with disabilities to evaluate their needs and preferences when using portable electronic devices, as well as the usability of these devices among users with disabilities. He is currently a student member of the Human Factors and Ergonomics Society.