Factors Affecting Seniors' Perceptions of Voice User Interfaces

by

Randall Ziman

December 2017

Presented to the

Division of Interactive Design and Information Architecture

University of Baltimore

In Partial Fulfillment

of the Requirements for the Degree of

Master of Science

Approved by: ______________________________

Dr. Greg Walsh, IDIA Program Director

____

Dr. Bridget Blodgett, Associate Professor

Abstract

This paper explores how seniors perceive Voice User Interfaces (VUIs) and the factors that shape those perceptions. An experiment was administered to 15 seniors (over age of 65), in which the participants searched for information using a traditional keyboard/mouse interface and an experimental voice/touch interface. Throughout the experiment, the participants were observed and interviewed by an investigator. An analysis of the data collected showed that seniors perceive meaningful differences between the two interfaces in terms of learnability, usability, ease of understanding and helpfulness. The data also shows that seniors' perceptions of the interfaces were most strongly influenced by familiarity, usability, habit, aversion to typing, and efficiency of voice input. These findings have implications for improving the design of VUIs for seniors and accommodating universal accessibility.

Acknowledgments

I would like to thank my advisor, Greg Walsh, for his support and encouragement to complete this research, as well as the rest of the UB faculty who helped me prepare for this work.

I would like to thank my family for their endless support and encouragement to pursue my master's degree. I would especially like to thank my Grandparents, Ann and Geoff Ziman who helped immensely by putting me in contact with many of this study's participants.

Above all else, I would like to thank my partner Janelle Steffen. She has been by my side for every step of this journey and I could not imagine reaching this point without her.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

We are currently experiencing a dramatic aging of the population. In the United States, life-expectancy has improved from 47.3 in 1900 to 78.9 in 2015 ("Health, United States, 2016 - Individual Charts and Tables," 2017). For some groups, like white women, the changes are even more dramatic – increasing from 48.7 in 1900 to 81.3 in 2015. As people are living longer, the number of seniors (people over the age of 65) is growing rapidly. Data from the US census (2017) shows that in 2016, there were 49.2 million people over the age of 65 and that they comprised 15.2% of the population By 2060, the number of people is expected to grow to 98.2 million, and comprise nearly 25% of the population (Bureau, 2017).

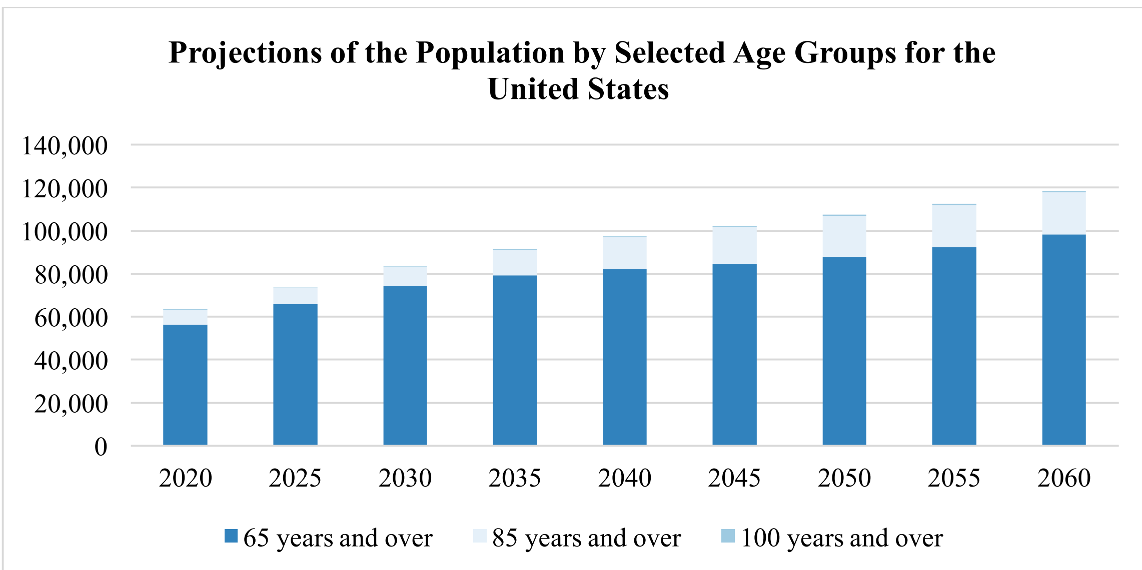

Figure 1. Data from 2014 National Population Projections Tables (Bureau, 2014)

In addition to living longer, seniors are finding technology an increasingly fundamental element of daily life. Not only has technology become an "integral component of work, education, communication and entertainment", it also supports seniors leading active, independent lives (S. Czaja & Chin Lee, 2009). Furthermore, "many new technologies are not optional — that is, older adults have no choice but to use them" (Fisk, Rogers, Charness, Czaja, & Sharit, 2009). For example, seniors may not have alternatives to using ATMs, automated phone systems and/or digital kiosks.

Together, the trends of increased life expectancy and technology use are driving the demand for products that are accessible and enjoyable for seniors. However, despite the fact that technology use among seniors is increasing, research shows that seniors still face unique usability challenges arising from changes in memory, attention, movement, and disability (Fisk et al., 2009).

Among the technologies available, Voice User Interfaces, VUIs, may hold potential for increasing usability for seniors. Many voice systems are efficient, intuitive and do not require the fine motor skills that older users can find challenging. However, research in this field is still limited and more work is required to identify the forces that shape user's perceptions of VUIs (Begany, Sa, & Yuan, 2016)

The goal of this paper is to gain a deeper understanding of how seniors interact with VUIs by identifying the major factors that affect user perceptions of VUIs. In addition to establishing new design principles for seniors, this work is likely to establish best practices that can support universal accessibility. Many of the age-related changes that affect seniors replicate the conditions that affect usability for a more general audience. In other words, be ensuring that designs are accessible for seniors, designers can increase accessibility for all users.

Chapter 2: Literature Review

The Diversity of Seniors

Seniors today are "healthier, more diverse and better educated than previous generations" (S. J. Czaja, 2005). One major factor contributing to the diversity of seniors is the increase in life expectancy ("WHO | Life expectancy," n.d.). As people live longer and the number of people over the age of 65 grows, the population of seniors becomes increasingly diverse. Consequently, the term 'senior' now collectively refers to 3 different generations: 1) the GI Generation born in the 1900s to the mid 1920s; 2) the Silent Generation born in the mid-to-late 1920s to the early-to-mid 1940s; 3) the Baby Boomers born in the early-to-mid 1940s to the mid 1960s). In *The Impact of Aging on Access to Technology*, Czaja recognizes these differences by distinguishing between seniors in general (over 65) and the "oldest old" (over 85) (S. J. Czaja, 2005). Building on the logical implications of generational differences, there is evidence that different generations of seniors use the internet for significantly different purposes (Street, NW, Washington, & Inquiries, 2010).

Beyond generational differences, multiple authors have recognized significance differences between individual seniors in attitude and aptitude regarding technology (Chisnell & Redish, 2004) (Gregor, Newell, & Zajicek, 2002) (Arch, 2008). Although trends in the use of technology by seniors can be identified, research shows that after accounting for other characteristics, age alone is not an ideal predictor of technology use after accounting for other characteristics (Charness & Root, 2009). A 2006 study of 1024 seniors found that abstract problem-solving ability, cultural knowledge, computer anxiety, computer efficacy, education, and ethnicity have a significant impact on technology use (S. J. Czaja et al., 2006). To illustrate the breadth of senior experiences, Morrell (2005) points out that a 66-year-old baby boomer who has been working in IT since the 1990s may experience technology in a dramatically different way than an 85-year-old, blue collar worker who has been retired for 20 years.

Seniors and Technology

Whereas seniors who used digital technology were once outliers, they are now a large and consistent group of technology users. Seniors' use of the broadband internet, smartphones, tablets and even social media have all grown dramatically between 2000 and 2016 (Anderson & Perrin, 2017b). See Table 1

Growth in percent of seniors adopting various technologies.

Table 1

Growth in percent of seniors adopting various technologies

	2000		2016	
	All Adults	Seniors (65+)	All Adults	Seniors (65+)
Internet	46%	12%	90%	67%
Home Broadband	1%	0%	73%	51%
Smartphone	35%	11%	77%	42%
Tablet	3%	1%	51%	32%
Social Media	21%	2%	69%	35%

Note (Anderson & Perrin, 2017b)

Though adoption rates for all measured forms of technology are increasing, the most dramatic of these shifts was in smartphone adoption, which nearly quadrupled between 2000 and 2016, raising to 42%. See However, these increases in adoption have not been made evenly by all seniors. Adoption rates are heavily skewed toward younger, better educated, and more affluent seniors (Anderson & Perrin, 2017a). See Table **2**

*Percent of seniors adopting smartphones by age*and Table 3

Percent of seniors adopting smartphones by household income.

Table 2

Percent of seniors adopting smartphones by age

	Smartphone Adoption
Total	42%
65-69	59%
70-74	49%
75-79	31%
80+	17

Note (Anderson & Perrin, 2017a)

Table 3

Percent of seniors adopting smartphones by household income

Household Income	Smartphone Adoption
< $30K	27%
$30K - <$50K	32%
$50K - <$750K	53%
$75K+	81%

Note (Anderson & Perrin, 2017a)

Age Related Changes

To understand seniors as technology users, it is vital to understand the age-related changes that mediate their experiences with technology. Fisk et al (2009) point out that though "age brings with it many capabilities," it also drives limitations (Fisk et al., 2009). Specifically, age-related changes leave seniors significantly more likely to exhibit physical and/or mental impairments that affect their use technology (S. J. Czaja, 2005). Fisk et al (2009), describe four categories of age related changes— sensation, memory, attention and movement— all of which can negatively affect seniors ability to use technology.

Sensation

Sensation is defined as "the awareness of simple properties of stimuli," including taste, smell, haptics, hearing, and vision (Fisk et al., 2009). Among these sensations, hearing and vision are the most widely researched in their connection to technology use. Over half of all men, and 30% of all women, over the age of 65 suffer from hearing loss that "hinders social interaction". Particularly, seniors have trouble with high-pitch tones (lower-pitch tones are generally unaffected by age related changes).

Visual impairments also correlate strongly with age. These visual impairments are primarily the result of cataracts, age-related macular degeneration, glaucoma, and diabetic retinopathy (Desai, Pratt, Lentzner, & Robinson, 2001). Though some of these issues can be treated, others cannot and the number of seniors affected by vision impairment continues to grow. According to the Center for Disease Control, in 2001, "approximately 1.8 million noninstitutionalized elderly report some difficulty with basic activities such as bathing, dressing, and walking around the house, in part because they are visually impaired" (Desai, et al., 2001). Visual limitations have also been identified as a contributing factor for lower word processing performance (S. Czaja & Chin Lee, 2009)

Memory

Dispelling the idea that memory necessarily deteriorates with age, Fisk et al (2009) clarify 1) there are multiple types of memory; and 2) each type of memory may be affected differently by age. Specifically, they identify four types of memory: working memory, semantic memory, prospective memory and procedural memory.

Working Memory Working memory is the capacity to "keep information active while we work on it or until we use it" (e.g.: remembering all of the digits of a phone number while we dial that number). Limitations on our working memory explain why we can remember a seven-digit phone number much easier than a twenty-digit number. Working memory is also subject to significant age-related declines. The implications of limited working memory are wide ranging and even extend to written and spoken language. The decline of working memory may cause more difficulty for seniors to make inferences, particularly when their working memory is already burdened by having to 'hold' the meaning of new words and concepts related to technology.

Semantic Memory Semantic memory is the capacity to access information that has been accrued through the course of one's life (eg: the meaning of a word or a historical fact). As opposed to working memory, semantic memory shows only "minimal decline" with age, although seniors may be slower to access semantic memory than younger adults.

Prospective Memory Prospective memory is the capacity for "remembering to do something in the future". Prospective memory can be time based (remembering to do something at a certain time, eg: at 2:00pm) or event based (remembering to do something after an event occurs eg: when the bell rings). Also, time based prospective memory is much more subject to age related decline than event based prospective memory, meaning that cues can be a valuable support for senior users of technology.

Procedural Memory Procedural memory is the capacity for remembering "how to perform activities" (eg: remembering how to ride a bike). Though older users may struggle to learn to perform *new* activities they maintain their ability to perform activities that have become successfully encoded over the course of their lifetime. This difference explains why an older user may be more comfortable using an arduous, but well-practiced method of interacting with technology than a new, more efficient method.

Attention

Although attention ("a capacity to process information") is closely entwined with working memory, Fisk et al (2009) also use this concept to point out differences in seniors' ability to selectively orient their attention. For example, seniors require more time to move their attention from one physical location to another. Furthermore, seniors tend to be more affected by 'distracting' stimuli such as flashing lights or even dense bodies of text.

Movement

Aging is strongly associated with diminished movement control, particularly for fine motor skills that are often required to manipulate technology (S. Czaja & Chin Lee, 2009). Numerous studies demonstrate challenges with precise movements such as "moving, clicking, fine positioning, and dragging". With practice, seniors are often able to improve their control, but continue to perform slower than younger users. Fisk et al

(2009) suggest that on average "older adults will be about 1.5 to 2 times slower than their younger counterparts".

Disabilities

Logically extending the implications of age based changes, it is unsurprising that seniors are significantly more likely to exhibit at least one disability. According to the US Census, 38.7% of seniors (over the age 65) reported at least one disability compared to only 20% percent of the general population (L. Larsen & He, 2014) (Office, n.d.). Also, many of these disabilities can affect the use of technology, with vision being and fine motor skills (like controlling a mouse) being particularly relevant. This situation creates the potential for seniors to be locked out of technology use and cut off from its benefits.

Benefits of Technology

Although age related changes may restrict and/or prevent seniors from using some digital technologies, appropriately designed technology can also be used to accommodate for these changes, increase independence, and improve quality of life (S. Czaja & Chin Lee, 2009).

Workforce

Nearly 20% of all seniors (over the age of 65) are still a part of the workforce in the United States ("Employment status of the civilian non-institutional population by age, sex, and race," 2017). In the Bureau of Labor Statistics most recent *Computer and Internet Use At Work Summary* ("Computer and Internet Use At Work Summary," 2005), 55.5% of all workers used a computer at work. Given the growth in computer adoption in the last 14 years, it is reasonable to expect that this rate has grown considerably. These seniors, who are using technology on the job, stand to gain the same benefits from technology as the rest of the workforce and adaptive technologies may even extend their ability to take part in the traditional labor force (S. Czaja & Chin Lee, 2009). Adaptive technologies such as screen readers, speech synthesizers, etc. may prevent workers facing age related changes and/or disabilities from shifting out of the workforce. Beyond the seniors who are using technology to improve the efficiency of their traditional functions, seniors also stand to benefit from increased opportunities for telecommuting. Seniors are more likely to have impaired mobility and/or receive some sort of in home care. Telecommuting also supports seniors' preference for more autonomy and flexible work schedules.

Home Environments

Technology is transforming home environments into a hub for information and services (S. Czaja & Chin Lee, 2009). Computers, mobile devices, and the internet now allow seniors to access information and vital services such as banking and shopping without leaving the home. This access is particularly important to seniors because they are more likely than younger people to face mobility issues that come from physical impairments, lack of transportation, inconvenience or fear of crime.

Stimulation and Social Connection

Several studies have shown that online communication, particularly email, is increasing social interaction among seniors. Seniors have also increased their use of social media in recent years (Anderson & Perrin, 2017b). Between 2008 and 2016, social media use grew 32% among seniors (Anderson & Perrin, 2017b). This social interaction has concrete benefits for seniors, especially those who live alone (S. Czaja & Chin Lee, 2009). For example, seniors who use the internet showed more positive attitudes toward aging, higher levels of perceived social support, and higher levels of connectivity with friends and relatives (J. Cody, Dunn, Hoppin, & Wendt, 1999).

Digital technology also increases access to educational activities that provide cognitive engagement and stimulation for seniors (S. Czaja & Chin Lee, 2009). In addition to fulfilling seniors growing interest in lifelong learning, this engagement and stimulation has been shown to facilitate "successful aging". Computers can even be used at home for cognitive rehabilitation and memory training (Chute & Bliss, 1994; Plude & Schwartz, 1996).

Healthcare

Technology also holds significant promise for increasing the quality and accessibility of healthcare for seniors (S. Czaja & Chin Lee, 2009). The same telecommuting benefits that allow seniors to participate with the workforce may allow them to communicate with healthcare providers. Research also shows that seniors are able to successfully conduct computer-based health assessments (Ellis, Joo, & Gross, 1991). Searching for health information is also one of the most common online activities for internet users of all ages, including seniors (Street et al., 2010). As baby boomers age, the data implies that over 80% of seniors will use the internet to search for health information.

Seniors and Voice User Interfaces

Introduction to Voice User Interfaces

As Voice User Interfaces (VUIs) become mainstream, they may offer a solution for connecting seniors with the benefits of technology while accounting for age based changes, and the challenges that those changes pose. VUIs, use voices as a medium for interacting with technology. This includes interfaces that use voice as an input, like Siri 'listening' to a user's question, or as an output, like an automated customer service system playing pre-recorded responses over a telephone line (Pearl, 2016). The 'voices' used as a medium can include natural human voices, recorded human voices or computer-generated voices (text to speech). Inputs, outputs, and levels of complexity vary greatly between VUIs. For example, some VUI systems will only be able to input specific spoken information like numbers or single address fields, others are more open ended; some VUI systems will output dynamically convert text responses to speech, whereas other will only play pre-recorded prompts. There is also some variance in how these features are described in the literature. VUIs are sometimes referred to Natural Language Interfaces (NLIs), Natural Language Systems (NLSs), Spoken Language Interfaces (SLIs), Spoken Language Systems (NLS) or Interactive Voice Response (IVRs) – although IVR is typically only used to describe telecommunications systems accessed via phone. Depending upon the focus of the research, these terms are often used to describe multi-modal systems that include voice inputs/outputs as well as Graphic User Interfaces (GUIs) and/or hardware such as buttons, or mouse devices.

VUI Benefits

VUIs offer seniors multiple advantages over traditional GUI/hardware interfaces by being hands-free, efficient, intuitive, and rich.

Hands-Free Perhaps the most apparent value of a voice based system is that they can be used without occupying the user's hands or eyes (Cohen, Giangola, & Balogh, 2004). Hands free capabilities may create opportunities for seniors who are required to use their hands for other purposes such as holding canes or walkers. Hands-free/eyes-free control will also benefit users who have limited ability to use of their hands/and or eyes, both of which are common limitations for older users. For example, a user with age-related macular degeneration may struggle to read a traditional GUI, but be very capable of navigating through audio (Corbett & Weber, 2016).

Efficient VUIs can be more efficient than hardware/GUIs for completing certain tasks. Researchers from the Stanford University and Baidu found that when using mobile phones, speech is a up to 3 times faster than input from touch based keyboards (Ruan, Wobbrock, Liou, Ng, & Landay, 2016). Considering that senior users typically take 1.5 times longer to complete tasks than younger users, this feature creates a significant advantage (Fisk et al., 2009). VUIs also allow users to eliminate touchpoints in their interactions. Rather than having to unlock devices, find apps, open apps and interact with apps, VUIs give users the opportunity to ask a question or give a command immediately (Cohen, 10). Eliminating these touchpoints reduces the necessary demands on working and procedural memory (Fisk et al., 2009).

Intuitive Humans have an innate propensity for language and learn the rules of conversation from the time they are young (Cohen et al., 2004; Pearl, 2016). VUIs allow users to interact with technology using the same strategies and mental models that they have already learned to communicate with other humans. By relying on pre-existing knowledge and strategies, VUIs become more usable than traditional alternatives because they can reduce the load on users' working memory, semantic memory and procedural memory. A senior using the traditional keyboard/GUI interface of a laptop, who wants to use Google to find the answer to a simple question would be required to remember how to: 1) open their browser; 2) navigate to google; 3) enter their query; 4) execute a search; 5) find an appropriate search result; and 6) open that search result. In comparison, the two-step process of using voice enabled Google Assistant is much simpler: 1) remember the wake phrase (e.g.: OK Google); and 2) ask a question using natural language.

Rich Humans often struggle to derive information from purely written interactions (Pearl, 2016). Purely written communications are often subject to greater interpretation than spoken interactions because speech can be mediated by tone, volume, intonation and speed to reinforce the indented meaning. For example, text messages and emails often leave the recipient questioning the author's intended tone in ways that spoken communication would not (e.g.: is the speaker truly upset or being sarcastic?) Considering that seniors are more susceptible to distraction when using GUIs, voice may be an ideal medium for communicating vital information such as alerts and/or safety warnings (Fisk et al., 2009).

VUI Challenges

For all the advantages that VUIs may offer to seniors, they also pose considerable usability challenges when they are opaque, inaccurate and/or vulnerable.

Opaque As an emerging technology, VUIs are inherently less familiar than the traditional GUIs that have dominated interactions for since the 1980s. Whereas, GUIs are designed to make hidden functionality visible to the user by showing menus, options, etc.; VUIs have fewer opportunities for implying their use to unacquainted users (Corbett & Weber, 2016; Yankelovich, 1996). Corbett and Webber (2016) illustrate how established best practices that support discoverability in GUIs break down when applied to VUIs, particularly mapping, constraints and affordances. Essentially, a VUI cannot passively indicate its own use because the interface is not discernable to users until it is in use. Therefore, designers cannot use analogies to imply use (mapping), limit inputs (constraints), or provide conceptual insight into how users should interact (affordances) (Corbett & Weber, 2016).

The inability of VUIs to imply their own use requires the users to make more assumptions about how the interface will work (Yankelovich, 1996). Yankelovich (1996) identifies two ways that these user assumptions can undermine user experiences, particularly how users learn and discover what the system is capable of executing:

1. Users will assume the system can understand more than it is actually capable of understanding
2. Users will be unaware of functionality that is available

Similarly, according to Karsenty (2002) ”spoken natural language interfaces are characterized by a high degree of opacity” which prevents users from making accurate mental models (Corbett & Weber, 2016).

Although seniors (some more than others) can learn to use unfamiliar technology and establish new mental models (particularly with support from training programs), this requirement increases the strain on seniors' working and procedural memory (S. Czaja & Chin Lee, 2009; Fisk et al., 2009). In other words, for seniors with diminished working and procedural memory, it may be more difficult to learn to user VUIs than it is to rely on established principles to operate GUIs. Also, learning these unfamiliar technologies may be further hindered by seniors' aversion to learning through trial-and-error (Leung et al., 2012).

Inaccurate In the words of Cohen (2004) “Every error degrades the usability of the system” and VUIs are uniquely prone to errors. First, the nature of an invisible interface makes it harder for users to recognize errors. Second, the technology that powers VUIs is subject to interference and distortion from the surrounding environment. Pearl (2016) points out that some speech recognition systems may have accuracy as great as 90%, but only when being used by an “adult male in a quiet room with a good microphone”. In the ‘real world’ accuracy will notably decrease. This aspect of VUI usability is particularly important as recognition accuracy has a significant effect on user satisfaction (Kamm, Walker, & Litman, n.d.).

Beyond the normal frustration of receiving inaccurate responses, seniors may be particularly vulnerable to erroneous responses. For example, a key process that makes VUIs function is establishing when the user has finished speaking (Pearl, 2016). Many VUIs determine the end of a user's input identifying an “end-of-speech timeout” i.e. a pause that is long enough to signify the end of the users input. If the pause is too short, the system may cut off the user before they have finished speaking. If the pause is too long, the user will wonder if the system accepted their input. Considering that seniors are typically slower to access semantic memory, and face more strain on their working memory, it stands to reason that seniors may need extended lengths for timeout (though there is paucity of research testing the necessary length of timeouts for seniors) (Fisk et al., 2009).

Seniors may also be more prone to erroneous responses because they are more likely to experience voice disorders (defined as "any time the voice does not work, perform, or sound as it normally should, or interferes with communication") (Roy, Stemple, Merrill, & Thomas, 2007). A 2007 study of 117 participants over the age of 65 showed that 47% of participants had faced a voice disorder at one point; and 29% faced a disorder at the time of the study. If seniors struggle with voice disorders and/or feel that they are likely to face voice disorders in the future, they may perceive less benefit in voice based interactions (Venkatesh, Morris, Davis, & Davis, 2003).

Vulnerable Using a VUI requires users to speak and/or receive information audibly, making them vulnerable to breaches of privacy. Depending on the content on that information, this could cause two types of privacy concerns. First, information being input/output is vulnerable to eavesdropping. For example, a VUI would not be well suited to helping users enter banking information in a public place because that information could be easily overheard. Second, VUIs are subject to security breaches from malicious actors. In other words, compromised VUIs may be used to covertly record surrounding conversations.

An AARP Public Policy Institute study of Internet users found that 75% of seniors are at least somewhat concerned with their privacy while "digitally connected" (Walters, 2017). With the advent of VUIs (and other devices), particularly as part of home appliances like Siri and Echo, seniors are living more of their life digitally connected (Anderson & Perrin, 2017b). Furthermore, 86% of seniors expressed concerns about their personal data being compromised (Walters, 2017). On the other hand, these concerns have not deterred the general trend of seniors adopting more technology in recent years (Anderson & Perrin, 2017b).

Technology Adoption by Seniors

Considering the strong opportunities and challenges that VUIs pose for seniors, the question remains: 'Will seniors actually choose to use VUI technology?'

Factors Affecting Adoption of Technology

There is a wide body of literature and a multitude of theories that seeks to understand and explain why some people adopt new technologies and others do not (Lee, Mehler, Reimer, & Coughlin, 2015). Two of the most prominent frameworks are 1) the Diffusion of Innovations Theory (DIT) and 2) the Technology Acceptance Model (TAM). In the Diffusion of Innovations Theory, Rogers identifies five characteristics that influence users' decision to adopt/reject new technology: relative advantage, compatibility, complexity, trialability (a user's ability to try the technology), and observability (Rogers, 2003). In the Technology Acceptance Model, Davis et al (Davis, Bagozzi, & Warshaw, 1989) proposed an alternate model of adoption/rejection: that users make their choices based on the perceive usefulness and ease of using new technology. The original TAM has been employed, applied, and modified by a variety of theorists (Venkatesh, Morris, Davis, & Davis, 2003). Notably, in *User Acceptance of Information Technology: Toward a Unified View*, Venkatesh et al (2003) expanded Davis's original TAM to include four determiners (performance expectancy, effort expectancy, social influence and facilitating conditions) and four moderators (gender, age, experience, voluntariness of use).

Seniors and Technology Acceptance

The underpinning of both the frameworks discussed above is that the benefits of the technology must outweigh the costs. However, seniors perceive the costs and benefits of technology differently than younger users (Chen & Chan, 2011). For example, Chen and Chan (2005, p. 200) point out that younger users perceive ease of use differently than older users. Whereas, younger users referred to the time necessary to complete each task (i.e. task efficiency), seniors focused on their ability to complete tasks successfully (i.e. task effectiveness). McCreadie and Tinker (2005) also point out that 'felt need' is important for seniors to adopt technologies. In other words, though seniors might be less interested in technology for its own sake, they are willing to use technology that can make their life easier (Chen & Chan, 2011).

Another factor that can affect seniors' assessment of technologies is that they typically have more trouble learning to use computers, less knowledge about computers, and less access to computers (S. Czaja & Chin Lee, 2009). A 2005 study by the Neilsen Normal Group, showed that websites were twice as difficult for older people to use than for younger people (S. Czaja & Chin Lee, 2009). Perhaps as a result of usability challenges, people over the age of 65 also tend to report lower levels of confidence and higher levels of anxiety when using computers, which is associated with lower interest in technology (Charness & Root, 2009).

Despite these trends, age alone may not be the best characteristic for understanding differences in perceived cost and benefits. Research from Czaja et al. (2006) shows that computer anxiety, fluid intelligence, and crystallized intelligence significantly affect seniors' use of technology. That same study showed that the "relationship between age and adoption of technology was mediated by cognitive abilities, computer self-efficacy, and computer anxiety".

The Power of System Experiences

In searching for a better predictor of technology, there is an increasing body of research comparing the various factors that may affect user perception and adoption of technology. According to Lee et al. (2015), this work can be grouped into three major categories (however, it is also important to note that the factors from each of these categories may affect one another):

1. Individual characteristics (e.g.: age, internet access, health, wellness, and functional capabilities)
2. Preconceptions and beliefs (e.g.: trust of technology, transference of perception from similar technology)
3. System experience (e.g.: did the user enjoy his/her experience)

To further understand the nature of these modifying influences, Lee et al. (2015) created series of hypotheses that describe each potential relationship between each factor and the resulting attitudes and perceptions toward technology and the compared (see Figure 2. Lee's research model)

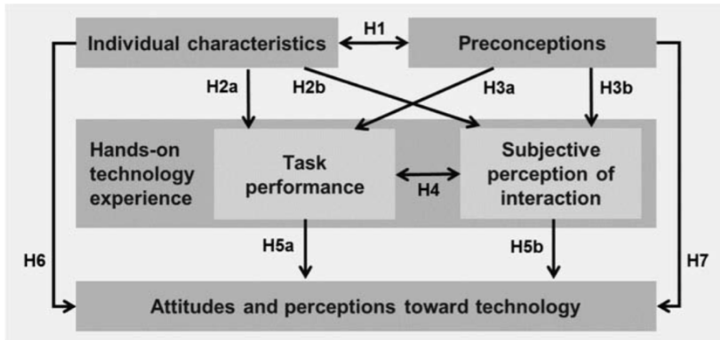

Figure 2. Lee's research model (Lee et al., 2015)

To compare the strength of these hypotheses, Lee et al. (Lee et al., 2015)analyzed the correlations between each contributing factor and the user's eventual attitude and perception of technology in three case studies related to in-vehicle technologies. Through this work, the team found little support for the hypothesis that individual characteristics (such as age, health, general preconception and task performance) affect the users resulting attitudes and preconceptions. However, their analysis did suggest a strong link between a user's perceptions of their hands-on system experience and their later attitudes and expectations. It is worth noting that the participants in Lee's study were all under the age of 69, with an average age of 43.62 so this study may fail to account for difference between seniors and younger users (Lee et al., 2015). On the other hand, this work does help explain why some seniors may adopt new technologies while others do not.

Measuring the Usability of VUIs

The importance of the hands-on system experience implies that the adoption of VUIs will likely be driven by the usability of specific systems rather than the overarching characteristics of voice technology in general. Unfortunately, measuring VUI usability has been a challenging endeavor for researchers. In 1999, researchers from AT&T created one of the first comprehensive frameworks for evaluating and comparing spoken language systems called PARADISE (PARAdigm for DIalogue System Evaluation) (Walker, Litman, Kamm, & Abella, 1998). Rather than focusing on singular metrics related to dialogue (eg: task completion time), PARADISE introduced a formula that estimated the system performance based on task completion rates *and* dialog costs. Although this system provided valuable information, it was primarily focused on evaluating the 'digital agent' rather than focusing on a user's experience with system.

In 2004, Dybkjær and Bernsen opened their paper on usability testing for VUIs by questioning Kamm's assumption that dialogue cost is a relevant driver of user satisfaction. After recognizing the difficulties and relative infancy of VUI usability testing the team provides a comprehensive review of evaluation practices for assessing the usability of systems with Voice User Interfaces. In their recommendations, they emphasize the importance of testing using realistic situations that do not prime the user on how to interact with the system. For conducting the actual usability tests, they recommend a template to support consistent and detailed evaluation. The template includes the following issues: what is being evaluated, the system part evaluated, type of evaluation, method(s) of evaluation, symptoms to look for, life cycle phase(s), importance of evaluation, difficulty of evaluation, cost of evaluation, and support tools. In relaying the nuances of evaluation, Dybkjær and Bernsen (2004) point out that "[VUIs] are very different from more traditional interactive systems whose usability aspects have been investigated for decades, such as systems controlled through graphical user interfaces involving screen, keyboard and mouse." To remedy this lack of research, they suggest 15 usability issues that should be considered for a comprehensive evaluation of usability in voice user interfaces: modality appropriateness; input recognition adequacy; naturalness of user speech; output voice quality; output phrasing adequacy; feedback adequacy; adequacy of dialogue initiative; naturalness of the dialogue structure; sufficiency of task and domain coverage; sufficiency of the system's reasoning capabilities; sufficiency of interaction guidance; error handling adequacy; number of interaction problems; and user satisfaction. In discussing user satisfaction, they point out that "much remains to be discovered about how the behavior of [VUI]s affect the satisfaction of their users" and recommend using questionnaires and interviews for gaining more information.

In line with Dybkjær and Bernsen's suggestions, Forbes-Riley and Litman (2011) used survey questions to assess user satisfaction in multiple studies considering the effectiveness of a spoken language tutoring system called ITSPOKE. After students interacted with a speech-based tutoring system, they were asked to rate their degree of agreement with 16 statements about the system's usability that focused on their subjective perceptions of likability, ease of use, effectiveness, etc. For each statement, the student's response was recorded on a five point Likert scale: almost always (5), often (4), sometimes (3), rarely (2), and almost never (1).

In *Assessment of Spoken Dialogue System Usability - What are We really Measuring?*, LB Larsen (2003) builds on Dybkjær's concerns about the lagging development of quality usability standards for VUIs. Larsen points out that despite the progress in objective measurements of usability such as success rates, recognition rates, and barge-ins, subjective measurements remain problematic. Although interviews and questionnaires are an ideal method for recording user perspectives, these methods are being employed in a mostly ad hoc manner and are rarely validated. Larsen points out the progress that has been made in developing a validated questionnaire by the Center for Communication Interface Research (CCIR) at Edinburgh University with British Telecom in the "Intelligent Dialogue Project" (the CCIR-BT) and Hone and Graham's Subjective Assessment of Speech System Interfaces (SASSI). Although, it is worth noting that SASSI was tested with 214 users interacting with 4 systems. The study suggested six main factors in users' perceptions of speech systems: system response accuracy, likeability, cognitive demand, annoyance, habitability and speed that "face validity, and a reasonable level of statistical reliability" (Hone & Graham, 2000)

In the years since Larsen's article was published, additional questionnaires have been developed, but no industry standard has been established (Lewis). In his review of the available resources for usability testing VUIs, Lewis recommends three questionnaires: 1) The Mean Opinion Scale (MOS); 2) Subjective Assessment of Speech System Interfaces (SASSI); 3) Speech User Interface Service Quality (SUISQ). Of these questionnaires, only one (SASSI) is designed to assess general speech system usability.

Rather than relying strictly on questionnaires for insight, Begany et al (2016) used variety of tools, including a quantitative content analysis to investigate subjective user experiences. This study explored differences in user perceptions when using a VUI/touch screen search system and a more traditional GUI/keyboard search system. During the course of the test, each user was given a variety of scenarios and asked to use one of the search systems (VUI or GUI) to find appropriate information. Information was recorded the users in a variety of ways: An entry questionnaire gathered demographic and other background information; a pre-task questionnaire elicited information about participants' knowledge of the topic; a post-task questionnaire elicited opinions about the particular search; a post-system questionnaire collected opinions about the specific system; and, an exit interview.

Current Study Contributions

Using a qualitative content analysis, Begany et al. (2016) determined several key factors that affect user's perception of VUIs: the users' familiarity with the system; the ease-of-use of the system; the speed of the system; the extent the user trusted the system; the user's comfort level with the system, 'fun factor' and novelty. The team also found that users perceived the VUIs differently that GUI/touch systems in terms of helpfulness, ease of learning, ease of use, difficulty and satisfaction. Although this information is valuable for helping designers and technologists design VUIs that will be perceived positively and primed for adoption, it was limited by its focus on younger users. Of the 48 participants studied, none were over the age of 65.

The primary focus of this research is to continue filling the current research gap regarding the perceptions of voice user interfaces identified by Begany by studying the perceptions of seniors i.e. users over the age of 65 (Begany et al., 2016). More specifically, the goals of this paper are to 1) gain a deeper understanding of how seniors interact with VUIs; 2) compare VUI interactions with GUI/touch screen systems; and 3) identify the major factors that affect user perceptions of VUIs. In addition to making a unique contribution to the existing literature, this work will provide insights that can improve experiences for seniors and promote universal accessibility.

By investigating how seniors perceive their experiences using VUIs, this paper strives to understand if and how VUIs are fulfilling seniors needs and expectations. Our hope is that this information can be used by designers and technologists to improve designs of existing VUIs and spur the innovation of new products that capitalize on these costs/benefits. This information may also inform our understanding of if and how mainstream seniors are likely to adopt VUIs. Furthermore, by comparing seniors experience using VUIs and GUIs this work has the potential to inform if and why users might be willing to abandon traditional systems in favor of emerging technologies.

Although this work will focus on seniors, the benefits of this work is likely to extend to the general population. Rather than thinking of seniors as having fundamentally different needs, many authors argue that promoting usability for seniors (or any group that is particularly vulnerability to issues of accessibility) will also improve usability for all users (Smith, 2013). Building on a blog post from Thomas Armstrong, Smith (2013) argues that all users exist on a "continuum of competence" and that even the most capable users are subject to "situational disabilities". Horton and Queensberry echo this statement in A Web for Everyone (Horton & Quesenbery, 2014). They note that despite the fact that not all users face mental or physical disabilities that make reading difficult, all people read in situations that make reading difficult, such as poor lighting, bright glare or using small screens. Therefore, Horton and Quesenbery argue, by designing to accommodate universal usability, designers and technologists can improve the experience for all users. Considering that seniors are disproportionately likely to face a disability (or negative results of age-related changes), understanding if and why seniors find a system usable can lead the way toward more usable technology for all users.

Chapter 3: Methods and Materials

To investigate the factors affecting seniors' perceptions of voice user interfaces (VUIs), an experiment was conducted that compared seniors' experiences using a multimodal voice and touch interface (VTI) to a traditional keyboard/mouse interface (KMI). Fifteen seniors (over the age of 65) were invited to search for information online using each system while being observed by an investigator. At multiple points through the testing session, the participants were interviewed about their experience. The observations and interview responses from each session were analyzed to identify trends in how seniors interact with VUIs; the differences between VTIs and KMIs; and identify the major factors that affect user perceptions of each system.

Experimental Design

The core of the experiment consisted of instructing participants to search for information using each interface and then discuss their experiences with the interface. To control for system order, eight of the participants used the KMI first and seven participants used the VTI first. To control for question order and combinations, the questions and question order were varied for each participant. We created a series of testing protocols that reflected these controls and participants were assigned to a testing protocol in roughly the order that they volunteered to participate. See Table 4
Testing protocols.

Recruiting

To recruit participants for this study, the investigator primarily relied on pre-existing social circles to distribute an email asking seniors to participate in a research study. In some cases, this message was sent directly to potential participants. In other cases, it was forwarded by an intermediary contact with a stronger relationship to the intended recipient. The message to seniors introduced the investigator as a graduate student from the University of Baltimore, broadly explained the purpose of the research as studying "how people use technology", and invited the potential participant to meet for a one hour session where they would use a laptop and an iPad to search for information and to talk about their experience using each one. The message also emphasized that 1) no prior experience was necessary, 2) the investigators were seeking participants of all skill levels, and 3) the investigators would provide any and all necessary technology.

Table 4

Testing protocols

Participant	Task 1	Task 2	Task 3	Task 4	Task 5	Task 6
1	K-F1	K-I1	K-E1	V-F2	V-I2	V-E2
2	K-F2	K-I2	K-E2	V-F3	V-I3	V-E3
3	K-F3	K-I3	K-E3	V-F4	V-I4	V-E4
4	K-F4	K-I4	K-E4	V-I	V-I1	V-E1
5	K-E1	K-I1	K-F1	V-E2	V-I2	V-F2
6	K-E2	K-I2	K-F2	V-E3	V-I3	V-F3
7	K-E3	K-I3	K-F3	V-E4	V-I4	V-F4
8	K-E4	K-I4	K-F4	V-E1	V-I1	V-F1
9	V-F2	V-I2	V-E2	K-F1	K-I1	K-E1
10	V-F3	V-I3	V-E3	K-F2	K-I2	K-E2
11	V-F4	V-I4	V-E4	K-F3	K-I3	K-E3
12	V-FI	V-I1	V-E1	K-F4	K-I4	K-E4
13	V-F2	V-I3	V-I4	K-F1	K-I2	K-E3
14	V-F3	V-I4	V-E2	K-F2	K-I3	K-E1
15	V-F4	V-I2	V-E3	K-F3	K-I1	K-E2

Note: The V prefix indicates that the voice system was used and the K prefix indicates that the keyboard system was used. The F, I and E represent a factual, interpretive and exploratory task respectively. The number represents which factual, interpretive or exploratory search task was used. See Search Tasks section below for a detailed description of factual, interpretive and exploratory tasks.

Participant Demographics

Fifteen seniors were recruited to participate in the study, 3 from the Denver area and 12 from the South San Francisco Bay Area. Nine participants were female and six were male, see Table 5

Gender of test participants. All participants were over the age of 65 and volunteered to participate without compensation. Although the participants were highly homogenous in terms of race and ethnicity (all participants identified as white and none identified as of Hispanic origin), they were much more distributed in terms of other demographics; including age, education, and income (as self-reported).

Table 5

Gender of test participants

Gender	Number of Participants
Female	9
Male	6

None of the participants were from the GI generation (born in the 1900s to the mid 1920s), but the sample included members of the silent generation (born in the mid-to-late 1920s to the early-to-mid 1940s) and the baby boomers (born in the early-to-mid 1940s to the mid 1960). Importantly, the sample also included seniors under and over the age of 85-- a line at which Czaja draws a distinction between 'seniors in general' and the 'oldest of the old' (S. J. Czaja, 2005). See Table 6

*Participants by age range*Table 6

Participants by age range

Age Range	Number of Participants
65-69	1
70-74	1
75-79	4
80-85	5
86+	3

Participants were also relatively distributed in terms of highest level of education and annual salary, although skewed toward the higher end of both measures. See Table 7 *Participants by highest level of education completed* and
Table **8**

Participants by salary.

Table 7

Participants by highest level of education completed

Highest level of education completed	Number of Participants
High school or equivalent	0
Vocational/technical school (2 year)	1
Some college	3
Bachelor's degree	4
Master's degree	4
Doctoral/Professional degree (MD, JD, etc.)	3

Table 8

Participants by salary

Annual salary at time of retirement	Number of Participants
Under $20,000	2
$20,000 - $34,999	1
$35,000-$49,999	1
$50,000 - $74,999	1
$75,000 - $99,999	3
Over 100,000	4
Would rather not say	3

Tools for Data Collection

When participants arrived for a test session, the experiment relied on five key resources to collect information: 1) informed consent form; 2) background questionnaire; 3) search tasks; 4) post system interview; 5) exit interview.

Consent Form

To ensure that participants were adequately informed and willing to participate, each person reviewed and signed an informed consent form. The consent form (see Appendix A) was derived from the University of Baltimore template and specifically covered the purpose, procedures, potential risks (which were minimal), and voluntary nature of the research. Participants were specifically asked to consent to the recording of their image and/or voice. This form was also used to relay the contact information for the investigators should any questions or concerns arise.

Background Questionnaire

The background questionnaire was used to document information that could be helpful in identifying trends as well as looking for gaps in the sample population. The questionnaire was formatted into 23 multiple-choice questions. For most questions, the participants selected the answer that best reflected their background. Some items specifically asked the participant to "check all that apply". Inspiration for the phrasing of questions was drawn from Survey Gizmo and Survey Monkey resources with advice for phrasing demographic questions and answers ("General Demographic Questions for Any Survey," n.d.; "How to Write Better Demographic Survey Questions (With Examples)," n.d.).

The first section of the questionnaire (questions 1-6) covered the participants' demographic information, including their: age, race, ethnicity, income, and education. The second section (questions 7-14) recorded information about the users experience with technology in general including the participants' self-assessed ability to use computers and smartphones/tablets. The third section (questions 15 – 23) addressed how long and how often participants have been using voice activated technology. See Appendix B for full list of questions.

Search Tasks

Although the participants completed a series of search tasks (tasks that required the participant to search for information), the information collected was not of primary concern to this study. Instead, the purpose of the search tasks was to ensure that the user was exposed to the interfaces in a realistic manner so that the other tools could record more accurate results. Search tasks are an ideal activity in this regard because the activity is highly familiar (the second most popular online activity for users of all generations), and the specifics of the task can be easily varied (Street et al., 2010).

Each search task was comprised of a short scenario and a related task that required the user to search for information. The structure of these tasks replicated the work of Begany's (2016) original study, which itself was inspired by Borlund's *The IIR evaluation model: a framework for evaluation of interactive information retrieval systems* (Begany et al., 2016). Bouland (2003) suggests that coupling search tasks with short scenarios "triggers and develops a simulated information need by allowing for user interpretations of the situation, leading to cognitively individual information need interpretations as in real life" and provides an experimental control "by being the same for all test persons".

To capture potential variation related to the type of search task, search tasks corresponded to each of Kim's categories of search tasks: factual tasks, interpretive tasks and exploratory tasks (Kim, 2006). Factual tasks require the user to find specific, concrete information that is not open to interpretation. Interpretive tasks require the user to make inferences and evaluations in the pursuit of information, while still looking for a specific answer or answers. Exploratory tasks are the most open-ended of the three and require the searcher to make significant judgments about whether information is helpful and correct or not.

For the purpose of this study, a total of twelve search tasks were used. Three of the questions came directly from the examples listed by Begany (Tasks F1, E1, I1) (Begany et al., 2016). Significant inspiration was also drawn from Wildemuth and Freund's (2012) work on exploratory search tasks. Listed below is an example from each category of search task. See Appendix C for a full list of tasks.

Factual

Scenario. Recently, you were debating about which state has the highest voter turnout in the last election. A friend of yours said it was Minnesota, but you thought that it was somewhere on the east coast.

Search Task. Find out which state had the highest voter turnout for the 2016 presidential election. Save the document(s) where you have found the information required.

Interpretive

Scenario. Recently, a friend of yours was talking about his finding a new school for his Grandson. Their family is trying to decide between sending him to a charter school, a public school or a private school. You have decided that you would like to learn more about the differences between each type of school.

Search Task. Find out about the advantages and disadvantages of each type of school (charter, public, private). Save the document(s) where you have found the information required.

Exploratory

Scenario. A friend of yours who just turned 90 was told by her doctor that getting more low-impact exercise will increase her fitness and help her avoid injuries. Your friend does not use the Internet and has asked you to help create an exercise program for her.

Search Task. Put together two thirty-minute low-impact exercise programs that she could alternate between during the week. Save the document(s) where you have found the information required.

Post System Interview

To establish a baseline measurement of each participants' experience, participants were interviewed after searching for tasks on each system (KMI or VTI). By measuring the participants responses before moving on to the second system researchers were able to establish a baseline of usability for each system without the results being affected by comparison to the alternate system. Each participant answered four post system questions and the responses were measured on a five-point scale though the phrasing of the responses was varied to best match each question. See Appendix D for a full list of questions.

Exit Interview

Building on the information from the post system interview, the exit interview asked the participants to compare the two search systems: the traditional keyboard system and the experimental voice system. The exit interview was comprised of eleven questions. The first eight questions asked the user to compare the two systems in terms of helpfulness, ease of learning, ease of use, quality of results and overall differences. Participants described which aspects of the interfaces they liked the most, those they liked the least, and if they would suggest any changes to the systems in the future. See Appendix E for a full list of questions.

Testing Environment

Drawing on the advantages of observing users in their natural environment, investigators chose to interview and observe users in their homes rather than a lab environment. These advantages include: convenience to the participant, seeing the participant in context, cost effectiveness (Rubin & Chisnell, 2008). Fortunately, all participants except for one were able to meet in their own homes; and that one participant was interviewed in the home of a close friend who was also a participant. To balance the home context with an environment that is conducive to testing, the participants were asked (prior to the session) to identify a workspace that would allow the investigator and the participant to sit together while looking at a computer.

Technology

Search Engine

For both the KMI and VTI search system, all searches were conducted using the Google search engine (accessed at https://www.google.com/) via the chrome browser.

Traditional Keyboard Mouse System

For the KMI search tasks, participants were asked to use a Toshiba laptop with an external Bluetooth mouse. Prior to the start of the experiment, the laptop would be turned on, connected to the internet (all participants were able to provide access to a home network) and logged in to a dedicated account in the Chrome browser. By using a new chrome account for each participant, the search history from previous participants did not affect later experiences.

Making use of the KMI interface required participants to place their cursor appropriately in a search field, type in search queries and press the enter button (or click the search button) to run the search. Participants were instructed and allowed to use search fields on the Google home page, the search engine results page, and/or typing directly into the address bar (See

Figure 3. Google home page/search engine. Users conducted iterative searches by clicking into a search field, editing the query and re-running the search (see

Figure *4.* Google search engine results page with suggested search queries)

Figure 4. Google search engine results page with suggested search queries. When participants found information that they deemed valuable, they were instructed to bookmark valuable information by using the mouse to click the star icon in the upper, right-hand of the address bar. See steps below:

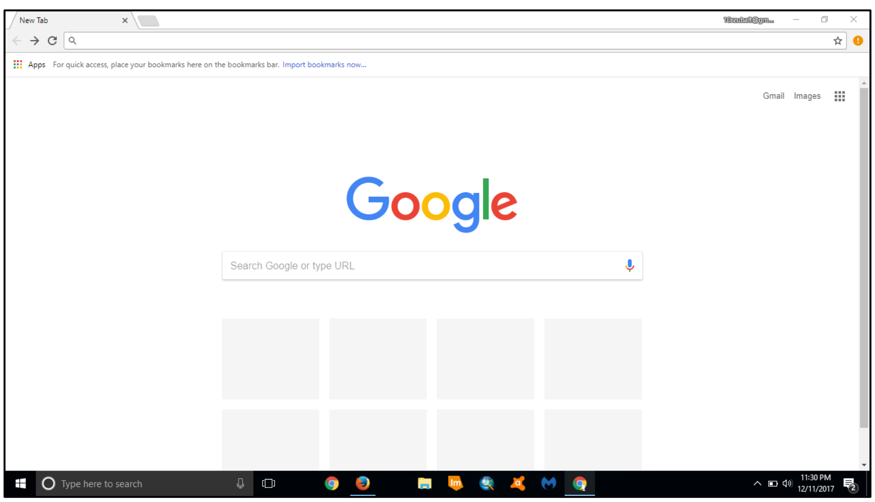

Figure 3. Google home page/search engine on computer with keyboard/mouse interface

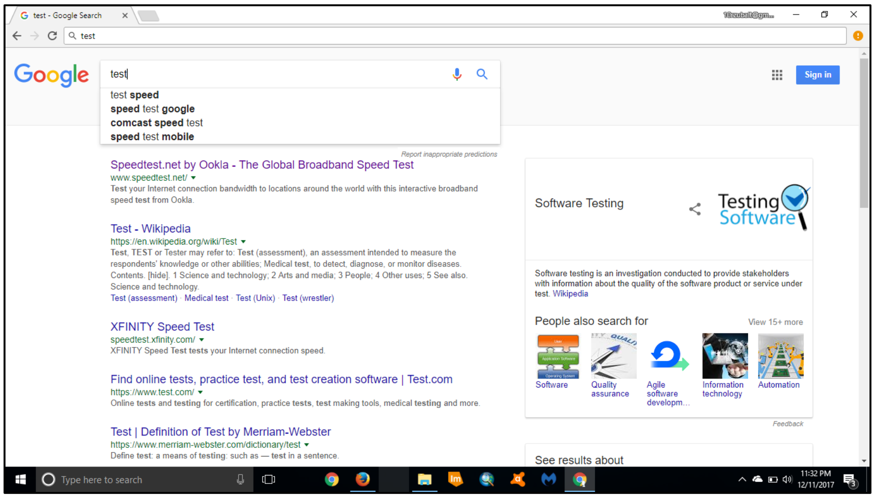

Figure 4. Google search engine results page with suggested search queries on computer with keyboard/mouse interface

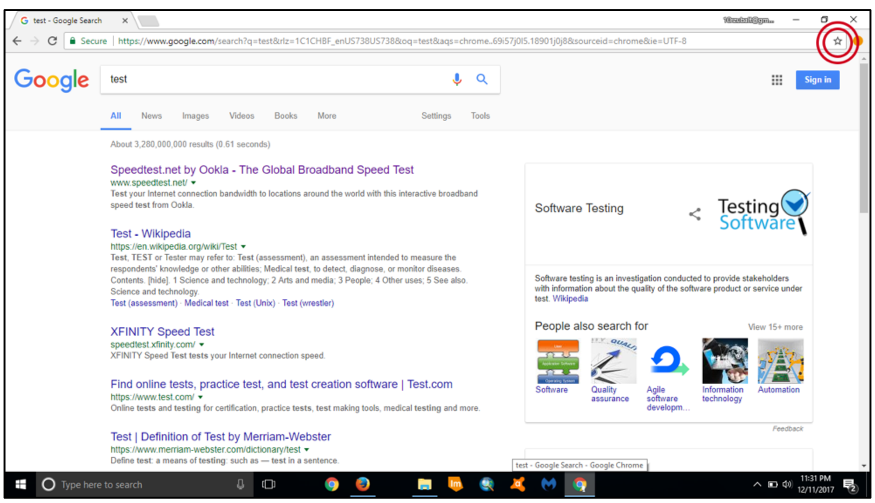

Figure 5. Google search engine results page with bookmark icon indicated on computer with keyboard/mouse interface

Experimental Voice Touch System

For the VTI search tasks, participants used an Apple iPad. Prior to the start of the experiment, the iPad was turned on, unlocked, and logged into a dedicated account in the Chrome browser (the same account that participant used for the KMI).

Making use of the VTI interface required participants to touch a microphone icon with their finger. Once users touched the microphone icon, the system chimed and the screen a voice input screen took over the interface. The voice input screen included instructions to "Speak now" until the user started talking. After a brief pause, the "Speak now" instruction would update to "Listening…". Once the system detected sound (i.e. voice input), the screen would display the user's voice input as it was being recorded. As the user continued to talk, the recorded input continued to update (the user's additional input provides more context for the system to recognize the word most likely intended by the user; e.g.: pair and pear may sound similar, but pear is a better fit in the search query "when does pear/pair season start?"). When the system detects a break in voice input, the search query (as recorded at that time) is run and the user is directed to a search engine results page.

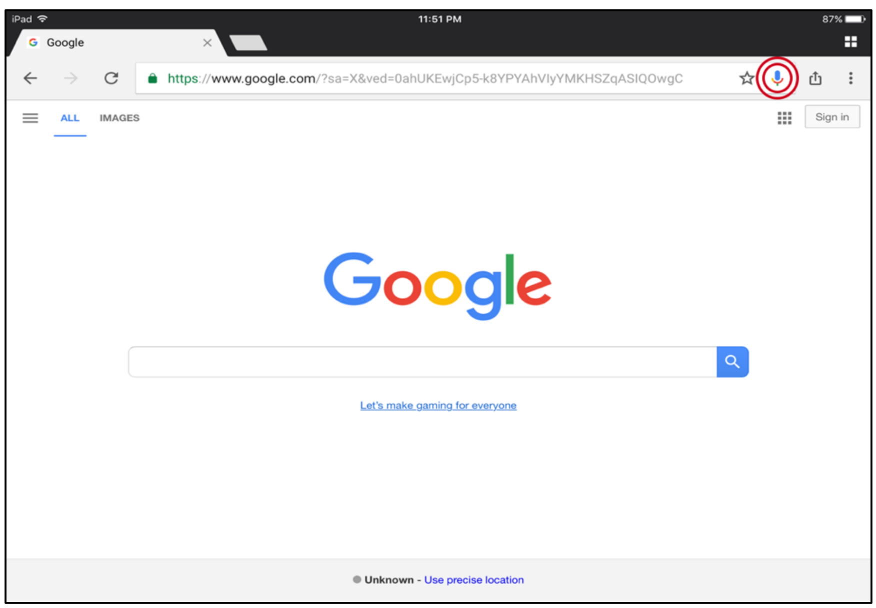

Figure 6. Google home page/search engine on iPad with voice interface

Figure 7. Google voice input screen on iPad with voice interface

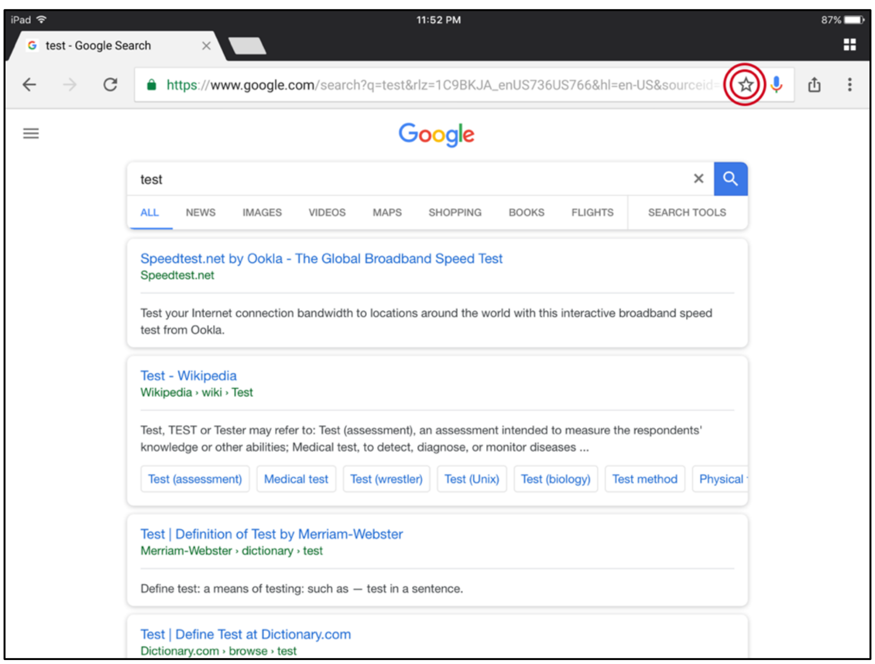

Figure 8. Google search engine results page with bookmark icon indicated iPad with voice interface

Testing Process

Overview

Each testing session included five key touch points: setup, introduction, search system 1, search system 2, and an exit interview.

Setup

Once the investigator arrived at the participants home, the participant was greeted to build a sense of rapport with the participant. The investigator asked to see the pre-identified workspace and set up that space with: 1) the search systems, 2) a personal computer for the investigator to take notes, and 3) a smartphone to record audio of the sessions.

Introduction

After the space was set up, the investigator guided the remainder of the session with the session script (see Appendix F for full script). The script started with an introduction that discussed the purpose of the study and a brief overview of the process. This section of the session included asking the participant to sign the informed consent form and complete the background questionnaire. After consent was given, the investigator started recording audio for the session.

Search System 1

To ensure that all participants started with similar knowledge about the technology a brief training was conducted. This training was limited to cover whichever search system the participant would use first (depending on their assignment in the testing protocol). In addition to explaining the steps for using the system, the investigator also demonstrated the process. After the demonstration, the user completed a practice search while encouraged to ask questions about the process. The search task varied between participants to match the same category as the users first search task (either factual, interpretive or exploratory).

After the practice task, the participant would complete three 'real' search tasks using their assigned system – one factual task, one interpretive task, and one exploratory task. For each task, the investigator would hand the participant an index card printed with the scenario and search task while reading the scenario and search task aloud. The participants pursued the search task for 7 minutes or until they felt that they had adequately completed the task. Completing the search task was determined by the participant’s own assessment. If the participant asked the investigator for approval the investigator instead asked if the user felt they had satisfied the task.

Once the participant had completed all three search tasks, the investigator moved into the post-system interview questions. The investigator read each question and the potential answers and recorded the participants answer. Participants often responded without using one of the potential answers in which case the investigator would push the participant to decide which option best reflected their answer.

Search System 2

When the participant was ready to move on to the second search system, the first system was removed from the workspace and replaced with the second search system. The format for the second search system was identical to the first: users trained to use the system, completed three search tasks, and participated in a post system interview.

Exit Interview

After the participant completed the second search system, the investigator moved into the exit interview. For each exit interview question, the investigator read the question aloud and then recorded the participants answers. In the event that the user offered an uncomprehensive answer, the investigator asked the participant to expand on the answer. Once all questions were adequately answered, the investigator thanked the participant for their time and concluded the testing session.

Content Analysis

After all of the interviews were conducted, the audio recordings of the exit interviews were transcribed and a qualitative content analysis was conducted. All transcripts were coded using an open coding approach to apply the codes established by Begany et al. (2016) – codes that were originally based on the themes that emerged from “a comprehensive read-through and evaluation of… 48 participants’ interview transcripts by each of the three investigators, the primary investigator and two doctoral students” See

Table 9

Begany's coding scheme for factors influencing users' perception for full list and description of codes. The codes were applied to the entirety of the exit interview transcript.

Table 9

Begany's coding scheme for factors influencing users' perception

Code	Factor	Description
FMLK	Familiarity	Familiarity with typing and the keyboard system itself
USBT	Usability	Spoken language system has good overall ease-of-use and usability
USBK	Usability	Keyboard system has good overall ease-of-use and usability
HBTK	Habit	User in the habit of typing/using keyboard
TRSK	Trust	User does not trust results from the spoken query; e.g. did the spoken system really understand?
NVFT	Novelty Factor	User enjoys the spoken language system because it is something new/new technology
SPDK	Speed	Speed of search results retrieval
CMFK	Comfort Level	User uncomfortable with voice system due to language barrier
FNFT	Fun Factor	The spoken language system is simply fun to use

Chapter 4: Results

Technology Experience and Self Efficacy of Participants

When asked to describe their technological expertise and experience on a background questionnaire, the participants clustered around 1-4 hours of computer and/or tablet use per day. Participants also tended to identify as 'intermediate' users. However, in terms of both experience and efficacy there was representation at both ends of the spectrums.

Table 10

Participants by technology expertise

Self-Rated Ability to use Technology	Number of Participants
Computer	
Novice	4
Intermediate	9
Expert	2
Smartphone or Tablet	
Novice	5
Intermediate	9
Expert	1

Table 11

Participants by technology use

Hours per day of use	Number of Participants
Computer	
Under 1 hour	5
2 – 4 hours	9
5 – 10 hours	1
Smartphone or Tablet	
Under 1 hour	7
2 – 4 hours	7

5 – 10 hours	0

Expanding from more generalized experiences with technology, participants were also asked to indicate their experience with and use of voice activated technology. Eleven out of 15 users indicated that they used a voice activated technology such as voice-to-text, Siri, Amazon Echo, Google Home, etc. Of the participants who used voice activated technologies, most participants started to use them 6 months to 3 years ago, but have only used them a few times a week or a few times a month. Only one user indicated using voice activated technologies every day. See Table 12
Participants by experience with voice activated technology and

Table **13**

Participants by frequency of voice activated technology use.

Table 12

Participants by experience with voice activated technology

Experience using Voice Activated Technology	Number of Participants
Never	3
Less than 6 months	2
6 to 12 months	4
1 to 3 years	4
3+ years	1

Table 13

Participants by frequency of voice activated technology use

Frequency of voice activated technology?	Number of Participants
Never	3
A few times a month	5
A few times a week	5
Every day	1
Multiple times a day	0

The number of participants who had experience using voice search declined slightly from the number of participants who used voice activated technology in general, implying that most users who use voice activated technology use that technology to search for information. See Table 14
Participants by frequency of voice search use. This idea is reinforced by the fact that searching for factual/ reference information is the most common online activity reported by participants and matches trends reported by Pew (Street et al., 2010, p. 201). See

Table **15**

Activities conducted online.

Table 14

Participants by frequency of voice search use

Use of Voice Search	Number of Participants
Never	6
A few times a month	4
A few times a week	3
Every day	1
Multiple times a day	0

Table 15

Activities conducted online

Activities conducted online	Number of Participants
News	10
Finance	5
Sports	1
Politics	7
Entertainment	6
Factual/Reference	12
Other	2

User Perceptions of Interfaces

Post System Interview Data

A comparison of post-system interview data shows that participants tended to generally look at both search systems favorable. By quantifying the participant's responses on a five-point scale (1 – very difficult, 2- difficult, 3-neutral, 4-easy, 5-very easy), the mean response is above 3 (neutral) for each question. See Table 16 *Average quantified post system response*.

Table 16

Average quantified post system response

Measure	Mean
How easy/difficult did you find learning to use the [system]?	
Keyboard/Mouse Interface	3.47
Voice/Touch Interface	3.8
How easy/difficult did you find using the [system] to search for information?	
Keyboard/Mouse Interface	3.73
Voice/Touch Interface	3.93
How easy/difficult did you find understanding the [system] interface?	
Keyboard/Mouse Interface	3.8
Voice/Touch Interface	3.33
How useful did you find the [system]?	
Keyboard/Mouse Interface	3.53
Voice/Touch Interface	4

Despite the generally favorable assessment participants made of both systems, there are some measures that return more pronounced differences between the two systems. When asked, "how easy/difficult did you find learning to use the [system]?", participants reviewing the voice system reported more favorable response (see Figure 9.

). Although drawing on quotes from the exit interviews (discussed in more detail below), the higher number of "neutral" responses regarding the keyboard system may be due to the fact that participants learned to use the keyboard system prior to the testing session (whereas the voice system truly did require learning on their part).

The second post-system interview question produced the greatest difference between the two systems. When participants were asked "how easy/difficult did you find using the [system] to search for information?", twelve participants described the voice system as "easy" or "very easy". On the other hand, only eight participants described the keyboard system as "easy" or "very easy". See

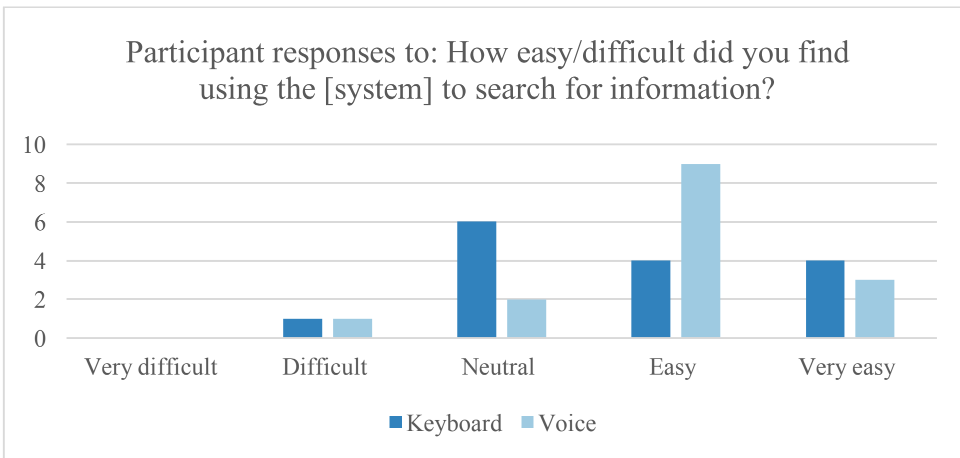

Figure *10.*

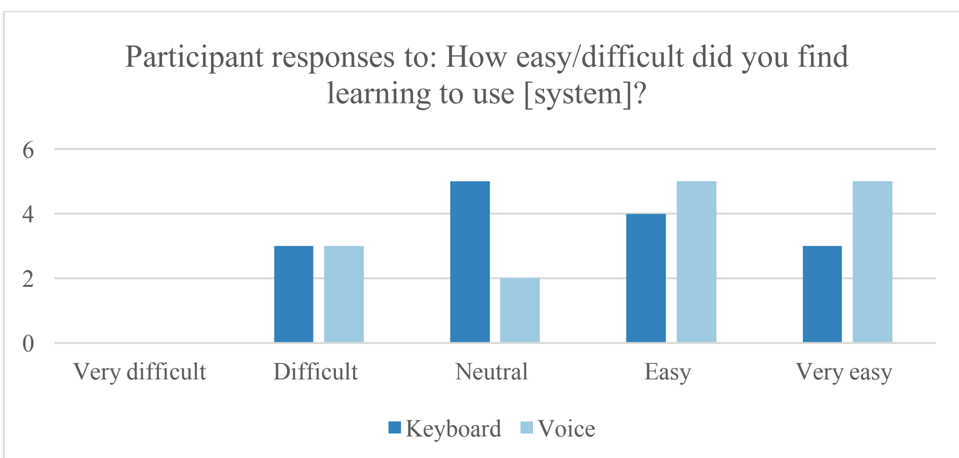

Figure 9. Comparison of interfaces by ease of learning

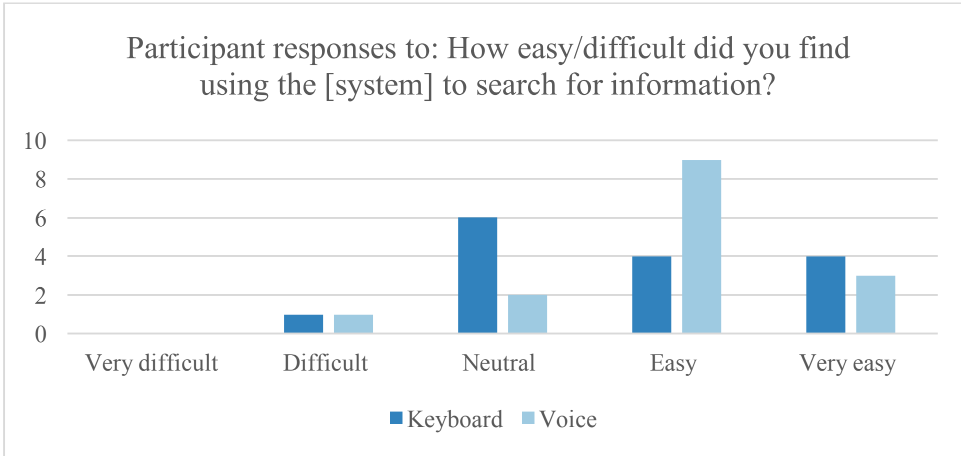

Figure 10. Comparison of interfaces by ease of use

Response regarding how easy or difficult participants found in understanding each system's interface are a bit more uniform, but it is worth noting that the ease of the voice system seems to top out. Whereas the voice system was most often rated "easy" to understand, the keyboard was most often rated "very easy". See

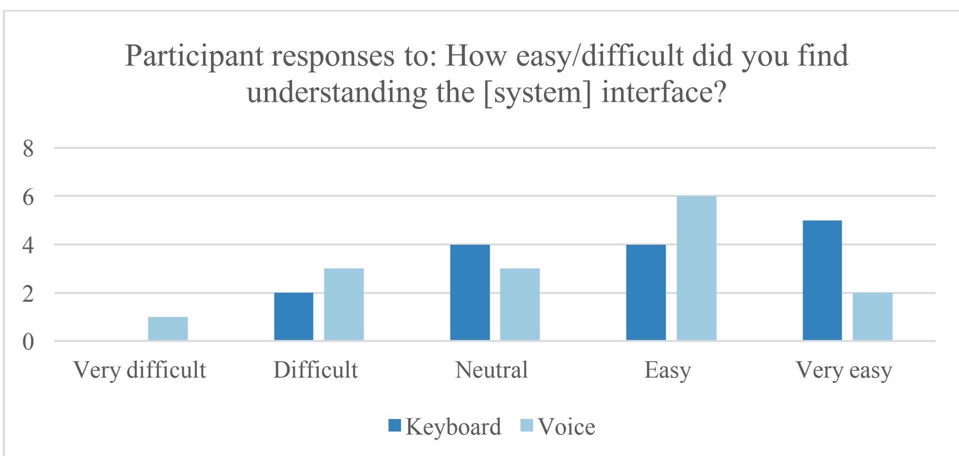

Figure *11.* .

In contrast to the other measures in the post system interview, the ratings in terms of usefulness hardly varied at all between the two interfaces. See Figure 12. . This uniformity may be explained by the fact that the use case was to search for information and the searches on both interfaces were provided by the same search engine/browser (a fact that was commented on by multiple participants during their exit interviews).

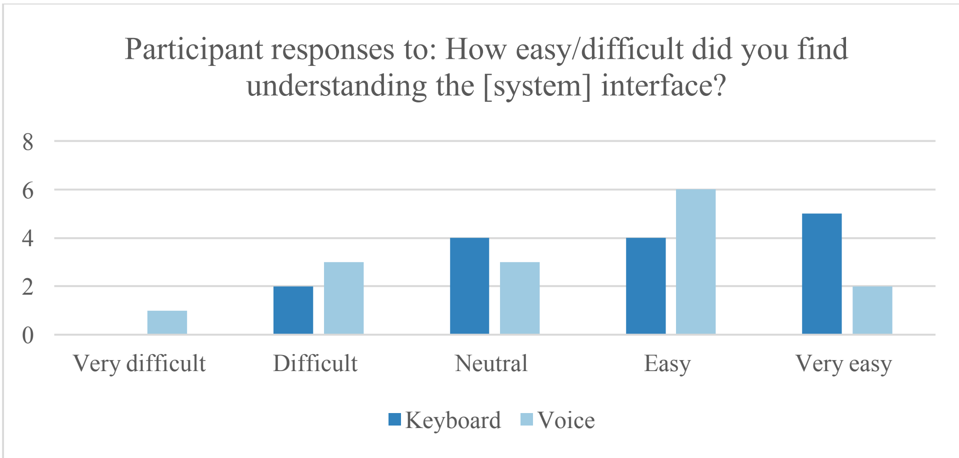

Figure 11. Comparison of interfaces by ease of understanding

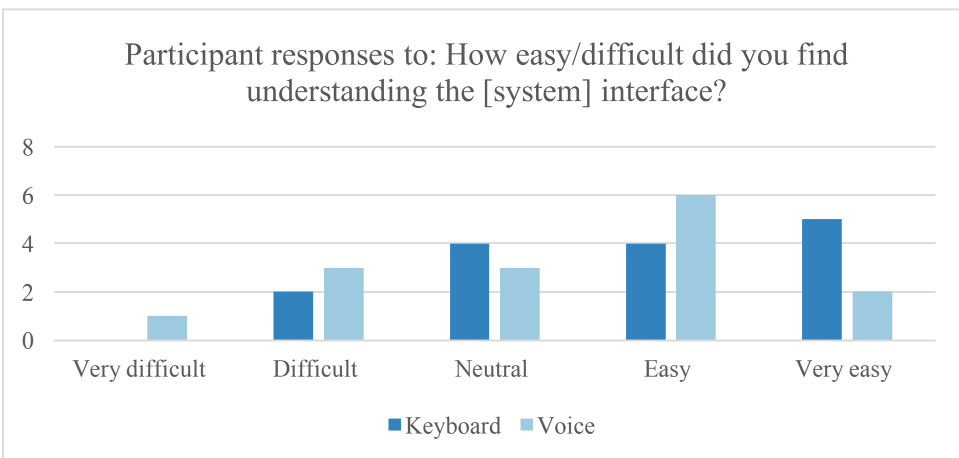

Figure 12. Comparison of interfaces by usefulness

Exit Interview Data

When asked to compare the keyboard and voice systems in the exit interviews, eight participants preferred the keyboard system overall compared with only five who preferred the voice system (two did not have a preference). Interestingly, the participants overall judgements do not seem to be an aggregation of the sub-factors relating to their experience. For example, participants reported that the voice system was easier to learn (by a margin of six) and easier to use (by a margin of five) while still preferring the keyboard system overall. See Figure 13. for full list of responses.

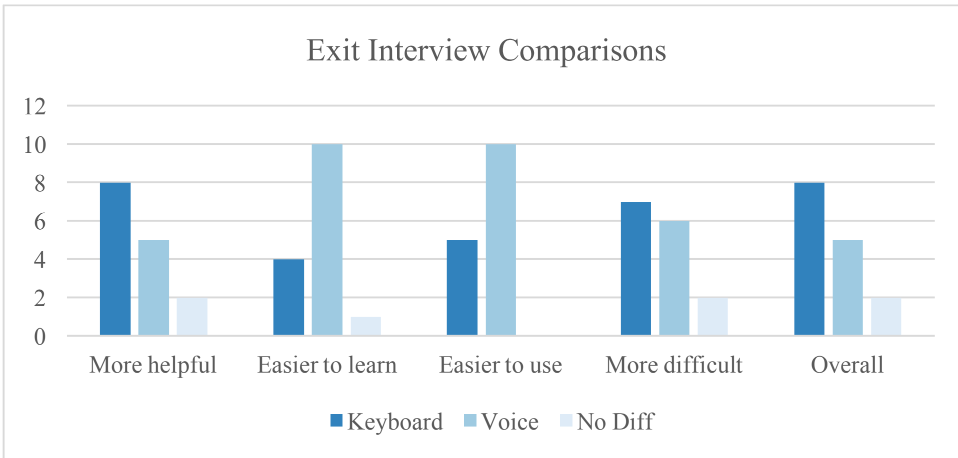

Figure 13. Comparison of Interfaces by exit interviews

Factors Affecting User Perceptions

Once the participants preferences were established, the content analysis of their exit interviews provided insight to the features that influenced their perceptions. Similar to Begany's findings, keyboard familiarity, keyboard usability, and voice usability ranked among the most influential factors (Begany et al., 2016). See Table 17 *Distribution of factors affecting users' perceptions*for full distribution of participants' responses.

Frequently Mentioned Factors

Although only 12 of the 15 participants explicitly mentioned their familiarity with the keyboard system, none of the participants noted its novelty. This lends extra significance to this factor as it can be assumed that every participant was familiar with the interface. However, it is worth noting that three users drew particular attention to the differences between the KMI used for testing and their own computers/laptops. In each of these three cases, the participants pointed out the differences between the Windows/PC

Table 17

Distribution of factors affecting users’ perceptions

Factor	Description	Participants
Familiarity (K)	Familiarity with typing and the keyboard system itself	12
Usability (V)	Spoken language system has good overall ease-of-use and usability	11
Habit (K)	User in the habit of typing/using keyboard	11
Usability (K)	Keyboard system has good overall ease-of-use and usability	8
Trust (K)	User does not trust results from the spoken query; e.g. did the spoken system really understand?	2
Novelty Factor (V)	User enjoys the spoken language system because it is something new/new technology	2
Speed (K)	Speed of search results retrieval (including query input and display of results)	0
Comfort Level (V)	User uncomfortable with voice system due to language barrier	0
Fun Factor (V)	The spoken language system is simply fun to use	0

system used for testing and their personal Apple computers (despite the fact that the Chrome browser is both available and essentially consistent across both platforms).

The second most commonly identified factor was the usability of the voice system. This is intriguing for because it contrasts with the overall preference for the keyboard system. Also, in the participants’ discussion of the voice usability they also identified multiple potential usability problems for the keyboard system. For example, one participant said, “I am not a typist”, presenting an interesting expression of generational perceptions about the role of typing. Other users mentioned the difficulties that they currently have with typing, some due to injury, some due to lack of experience. One user even discussed that he expected to lose his ability to type due to age related changes, at which point he would only be able to use a voice based system. Furthermore, multiple users pointed out that the voice based system removed the potential for spelling errors and typos, which seemed to be of significant concern to the seniors interviewed.

On the other hand, multiple users highlighted the advantages of the keyboards usability. The keyboard system was described as better for users to take their time entering a query and modify that query. In the words of one user, "when you're using the voice system, you really have to organize your thoughts before you commit to asking for information... not just organize your thoughts but organize them in such a way that you can verbalize them". This was particularly problematic for that user as she said "I think through my fingers.

As noted earlier, the habit of keyboard use stood out as a notable factor affecting users' perceptions. When users did explicitly mention their habits of using the keyboard system, they made it clear that they used these interfaces regularly and for a long time. For example, one user points out that she has been using a KMI regularly for "30 years".

Infrequently Mentioned Factors

Neither voice novelty, nor keyboard trust were mentioned frequently by participants of this study. Each factor was mentioned by only two participants and perhaps reflect a changing perception of VUIs and their capabilities since the time of Begany's original study (Begany et al., 2016). The users who mentioned trust were also observed to have more issues with VUI time outs, which seems to have been interpreted by the participants as the system "not understanding" them. The lack of novelty codes also corresponds to the participants familiarity with and use of voice activated technology. In the demographic tracker, 11 of the 15 participants recorded having used voice activated technology in the past. In other words, the interface only had the potential to be novel for 4 participants and 2 of them did mention this fact.

Unmentioned Factors

Out of the nine codes provided by Begany (2016), three of them went completely unmentioned by the sample participants. These codes were: keyboard comfort, keyboard speed, and fun factor. Despite the code's name, Begany describes keyboard comfort as "Comfort Level (user uncomfortable with voice system due to language barrier)". None of the seniors interviewed mentioned or demonstrated any sort of language barrier, so this code is essentially not applicable. Fun factor and keyboard speed, however, do have the potential to apply to the senior population, they just were not demonstrated. This is not to say that the seniors interviewed were not impressed with the system, but the content of their comments were better categorized as voice novelty and voice usability rather than fun factor. In terms of speed, rather than commenting on the speed of the keyboard system, many participants specifically mentioned the speed of the voice system. Multiple participants felt that the voice input was faster than typing and that it even returned results faster. These perceptions also match the investigators observations: that seniors were consistently able to voice input faster than they were able to type.

Additional Factors

While conducting the content analysis, two additional themes emerged among senior users: aversion to typing; and efficiency of voice input. In a second review of the interview transcripts, eight participants demonstrated concerns about typing and seven participants noted the speed and efficiency of voice input. As mentioned above, the concerns about typing arose due to fear of typos and spelling errors, physical constraints, and/or the effort required. Both of these factors are clearly demonstrated in one user's discussion of the voice system "I make less mistakes… it's just taking me less time to be able to speak in it rather than to type into it… I feel it's more efficient to be able to talk into it."

Table 18

Distribution of factors affecting users' perceptions

Code	Factor	Description	# of Participants Coded for Factor
AVRT	Aversion	Concerns about typing/input via the keyboard	8

EFFT	Efficiency	Efficiency/speed of voice input	7

Investigating the Effect of Self-Efficacy

In Cazja et al.'s 2006 paper, the team found that self-efficacy was a major factor in predicting the use of technology in general. Hypothesizing that self-efficacy may have a similar role in affecting users' perceptions and evaluations of various technologies the participants were segmented by self-efficacy. By comparing the results of the four users who self-rated their computer ability as 'novice' versus the eleven who rated 'intermediate' or 'advanced', investigators found series of sharp divides:

- 75% of self-rated computer novices preferred the voice interface compared to 11.11% of more advanced participants
- 100% of self-rated computer novices commented on the usability of the voice system, compared with 63% of more advanced participants
- 75% of self-rated computer novices commented on the efficiency/speed of voice system, compared with 36% of more advanced participants

This divide also held true when comparing the five users who self-rated as novice in smartphone/tablet ability with the ten users who self-rated as intermediate/expert:

- 100% of self-rated smartphone/tablet novices commented on the usability of the voice system, compared with 60% of more advanced participants
- 80% of novices commented on the efficiency/speed of voice system, compared with 30% of more advanced participants
- 60% of self-rated smartphone/tablet novices also preferred the voice interface compared to 22.22% of more advanced participants

Observations

Through the course of the interview session the investigator was able to identify common behaviors exhibited by the seniors interviewed. The three most prevalent observations were: 1) time out errors; 2) use of visual cues and supports; 3) distraction; and 4) mouse errors.

Time Out Errors

Every participant interviewed experienced at least one time out error. For the most part, these errors occurred when the users where pausing mid-thought while trying to construct their search query. This error was exacerbated by the fact that participants frequently looked up and/or away from the device when thinking, thereby missing any visual cues that an error had occurred. By the time the user looked back at the interface, either a search had been executed with only a segment of the users intended query; or the input had failed and none of the speech was input. Multiple participants expressed frustration at this point, although interestingly, participants did not seem to express as much frustration when struggling with the voice input as they did when struggling with the keyboard input (e.g. typos).

Visual Cues

Participants consistently relied on suggested and related searches when completing their search tasks. Even the users who did not realize that these features were 'click-able' used them as a prompt for modifying their searches. These features were used more often on the keyboard interface because users could interact with them as they input their search query. In their think aloud, participants often noted the existence of these search supports and referenced them when modifying searches. When using the voice interface, these features were used less because they were not visible until after the user had entered a query. Furthermore, if a user did want to rely on a suggested search while using voice input, the participant would have to remember the phrase they wanted to enter because the suggested search feature was not visible at the same time as the voice input screen.

Distraction

The seniors interviewed demonstrated a high potential for being sidetracked, particularly by invasive visual displays like pop-ups. When confronted with these displays, participants struggled to find buttons to close them. In some of the more extreme situations, participants decided to close and re-open the browser because they could not find a way back to their search results page. Furthermore, some participants seemed unable to differentiate between paid advertisements and authentic search results. This confusion sometimes led to users clicking through multiple pages of ads before realizing that they were only tangentially related to the information they were seeking.

Mouse Errors

Multiple users, and three in particular, demonstrated extreme difficulties with the external mouse. Specifically, they struggled with the nuances of the mouse interactions, ie: understanding when clicking the mouse would select an entire field (clicking into the address bar), vs when it would place the cursor (the search field). These users also struggled to understand when their cursor had been placed and when it had been placed and when it had not. After these participants had scanned the search results page, they often had trouble placing their cursor appropriately so that they could modify their search query. Lastly, some users had trouble with the fine motor skills required to move the mouse. Despite the fact that some users struggled to tap icons with a finger (particularly the microphone icon and the bookmark icon), it was more common that users struggled to click with the mouse especially when attempting to click small UI elements like the bookmark icon.

Chapter 5: Discussion

Differences in Perception of Interfaces

These results show that seniors perceived the two interfaces differently in terms of learnability, usability, ease of understanding and helpfulness. More specifically, the seniors interviewed perceived the VTI as more learnable and usable but the KMI as easier to understand and more helpful. The participants also ultimately preferred the KMI overall.

The finding that seniors prefer the VTI in terms of learnability can be directly related to the work of Cohen, Giangoal & Balogh (2004), and Pearl (2016). Each of these authors (as well as many others) discuss how VUIs are more intuitive and leverage a user's pre-existing speaking skills without requiring the user to learn new processes (thereby reducing the amount of new information to learn) . For example, one user went so far as to point out that it took her years to learn to type proficiently, but she was able to learn how to enter voice input in a matter of minutes.

The seniors' perception that the VTI was more difficult to fully understand also ties to established concepts in the field of VUI usability, particularly the work of Corbett & Webber (2016), Karsenty (2002), and Yankelovich (1996). These authors discussed the challenges arising from the opaque nature of VUIs especially that: 1) VUIs are less able to indicate their own use; and 2) users are less likely to make appropriate assumptions about how the system functions (Corbett & Weber, 2016; Karsenty, 2002; Yankelovich, 1996). It is also worth noting that seniors' observed behavior further confirms the challenges of understanding VUIs. For example, they struggled to understand what terms were used to execute a search compared to what they thought they input (either because the search omitted intended terms or included unintended terms).

The results of this study also provide insight about the comparative importance that different aspects of each interface play in forming an overall impression. For example, despite the fact that users found voice easier to learn and easier to use, they still preferred the KMI overall. This implies that the factors where the KMI was higher rated are likely more important than factors where VTI is higher rated. In this study, the participants rated the KMI higher in terms of ease of understanding and helpfulness. The KMI was also rated higher than the VUI in terms of providing better searches/results, but nearly twice as many participants found that there was no difference between the systems' quality of searches and results. On the other hand, the VUI scored better in terms of learnability and usability. Therefore, we can conclude that ease of understanding and helpfulness are more important than learnability and ease of use in seniors' perceptions of VUIs – a conclusion that builds on the findings from Chen and Chan (2011) that seniors prefer task effectiveness over task efficiency.

Differences Between Seniors and Younger Users

We can further understand the factors that affect seniors' perceptions of VUIs by comparing their feedback to the statements from younger users recorded by Begany et al. (Begany et al., 2016). These results show that the most commonly mentioned (i.e. most important) factors affecting seniors' perceptions of VUIS were: keyboard familiarity, voice usability, keyboard habit, aversion to typing, voice efficiency, keyboard trust, and voice novelty (in that order). Younger users most commonly mentioned: familiarity, usability (textual interface), usability (spoken language interface), voice novelty factor, keyboard speed, voice fun factor, keyboard comfort level, keyboard habit, and keyboard trust (in that order). See Table 19
Comparison of code distribution for a comparison of seniors' vs younger users.

As demonstrated in the table mentioned above, there are some factors important for both seniors and younger users, particularly the familiarity of the KMI, the usability of the VTI, and to a lesser extent, the usability of the KMI. For seniors, the importance of familiarity likely relates to the age-based changes to memory that are well documented by Fisk et al. (2009). Fisk et al. (2009) found that procedural memory (remembering how to complete routines and processes) is maintained in old age, whereas working memory (required for learning new processes) diminishes. This preference for relying on procedural versus working memory may explain why voice novelty was mentioned less often than other factors. This is perhaps even better explained by McCreadie and Tinker's

(2005) finding that 'felt need' (rather than just novelty) is important for seniors to adopt technologies.

Table 19

Comparison of code distributions

Factor	Seniors coded for factor	Younger users coded for factor
Familiarity (Keyboard)	80 %	79.17 %
Usability (Voice)	73.34 %	72.92 %
Habit (Keyboard)	73.34 %	14.59 %
Usability (Keyboard)	53.34 %	75 %
Aversion (Keyboard)	53.34 %	n/a
Efficiency (Voice)	46.67 %	n/a
Trust (Keyboard)	13.34 %	2.09 %
Novelty factor (Voice)	13.34 %	45.84 %
Speed (Keyboard)	0 %	45.84 %
Comfort level (Keyboard)	0 %	29.17 %
Fun factor (Voice)	0 %	37.5 %

Note. From (Begany et al., 2016) Aversion and efficiency are not applicable to the finding for younger users because those themes were not used in Begany et all's coding schemes.

There is one factor that seems to be dramatically more important for seniors than younger users: the habit of KMI use. Again, this could likely be explained by seniors' diminished working memory versus maintained procedural memory. Another option could be that habits are more meaningful to seniors because those habits have the potential to be more deeply ingrained than for younger users. For example, one senior mentioned that she had been using a KMI every day for 30 years. Considering that 80% of the younger users interviewed by Begany et al. (2016) were under the age of 35 (and would not have had the time to build up these kinds of long-lasting habits), it stands to reason that habitual use would be less applicable to their decision-making process.

Trust was also mentioned more often by seniors than younger users but was still one of the factors with the least mentions. Only 2 of 15 participants indicated trust issues with the VTI and they frequently did so in the midst of experiencing time-out problems. Therefore, trust is likely a less universal concern for seniors and perhaps better attributed to usability problems with the VTI rather than inherent trust issues. Another possible explanation is that VUIs have become much more prevalent in the years since Beganys' study. Eleven of the 15 participants used VUI prior to this study; and If users are familiar with VUIs and have seen the technology used successfully, their skepticism of voice input/trust that only the keyboard can correctly input information may no longer affect their perceptions.

There are also multiple factors that seem to be less important for seniors than younger users: novelty of the VTI, fun factor of the VTI, and the speed of the KMI. (Note: comfort level of the KMI is omitted from this list because none of the seniors interviewed faced a language barrier).

As discussed above, novelty may be less of a concern for seniors because of age related changes to working versus procedural memory and/or the importance of 'felt need' . Similar to the importance of trust, novelty may be less important due to the increased prevalence of voice technology. In other words, it could be that novelty is a factor that affects users' perceptions, but that this interface simply is no longer novel. The issue of felt need may also explain why fun-factor is less of a concern, but this could also reflect the participants disinterest with the search tasks. Despite the advantages of the standardized search tasks (discussed above), this approach did not allow the participants to experiment with the devices in an open-ended manner, which could have allowed each user to act in the way that they found the most interesting and fun.

Although the speed of the KMI was noted by younger users, it was universally dismissed by seniors. Seniors actually preferred and routinely praised the speed of the VTI. This difference could be the result of the challenges seniors reported and were observed to face when typing. Many of the seniors interviewed were not able to type fluidly, and/or were resigned to hunting and pecking with two fingers. This was especially the older seniors (75+). When comparing the KMI and VTI interfaces, she commented "I am not a typist" – a statement that would not likely be echoed by younger users and provides some intriguing insight into the generational differences and expectations about using technology. These typing difficulties were further complicated by challenges with fine motor skills. One of the participants, who self-rated as an expert computer user and had previously worked in highly technical fields, faced challenges with typing due to a shoulder injury. Another participant, who preferred the KMI was still adamant about the value of the VTI because "at some point, I won't be able to use a keyboard… this auditory staff is voice-actuated is probably a good, survival technique".

Interface Preferences and Self-Efficacy

Our understanding of the factors that shape seniors' perceptions of interfaces can also be enriched by considering the results in terms of self-efficacy. In the results of this study, participants who self-rated as their ability as 'novice' demonstrated a much higher preference for the VTI. These novice participants were also much more likely to comment on the usability and efficiency/speed of the voice system. Together, these findings imply that VTIs may be more accessible for less advanced/experienced users because of their usability and efficiency; and prior work in the field provides and explanation for why this may be true. Fisk et al. (2009) found that although older people were eager to learn how to use technologies they also perceived that they might have difficulty learning to use the systems and that they would require more time to learn than would younger people (S. J. Czaja et al., 2006). If seniors already know how to use a voice system from their use of voice in their everyday life, this might change their judgements about the perceived difficulty of learning the system. Also, other studies (e.g., Czaja & Sharit, 1998; Ellis & Allaire, 1999; Tacken, Marcellini, Mollenkopf, Ruoppila, & Szeman, 2005) have found that older adults expressed less comfort in using technology and less confidence in their ability to successfully use these systems. If seniors are confident in their ability to use a voice system from their use of voice in their everyday life, this might change their judgements about the perceived difficulty of learning the system

Improving Design and Usability for Seniors

Viability of VUIs for Seniors

The results of this study show that VUIs are a viable option for seniors. Despite the fact that seniors are often regarded as opposed to adopting new technologies, these findings build on the body of work showing that seniors can and will use voice based interactions. These findings seem to be particularly significant for working with novice users who are more likely than intermediate and advanced users to prefer a voice based interface. This implies that particularly when designers are working on untraditional or innovative interfaces for seniors, voice interactions should be considered as a viable option for lowering the perceived ease of use.

Controls for Executing Searches

The primary usability challenge for seniors using the VTI was time-outs. Either the participants were not able to fully input their request; or they accidentally inserted unintended words into their query. The observed seniors did tend to have more difficulties with omitted input than accidental input. Furthermore, the seniors were able to recognize patterns in their accidental input. Even though the seniors displayed signs of frustration about accidental input, they did realize that the connection between the words that they were saying and the words that were searched. What was more challenging for seniors was the disconnect between what they thought that they input and what was actually input. Another interaction that was challenging for the participants was that they could not pause to gather their thoughts. Participants would routinely begin a search query and stop half way through to think about how they wanted to finish their statement; and pausing would trigger a search that did not match the participant's intended query. This was especially problematic because the seniors frequently looked up and away from the display when thinking about their query, so they would often fail to notice that the search had omitted key words or phrases.

Seniors may benefit significantly from a setting that prevents searches from automatically triggering after a pause/timeout. This interaction pattern is already prevalent voice interactions for sending text messages. For example, when sending a text message through Google Assistant, the interface will ask the user to input the message and then ask the user to confirm or edit before sending the message. See Figure 14. Example text interaction using Google Assistant for an example of this interaction.

User: OK Google, send text message to Greg

Google Assistant: Text Greg, sure, what's the message?

User: Hello, this is Randall

Google Assistant: Got it, you want to send it or change it?

User: Send it

Google Assistant: OK, message sent

Figure 14. Example text interaction using Google Assistant

Applying this model to voice searches, the query could be held until the user confirms their query; either by giving a voice command or pressing a button. See Figure 15. Potential voice search interaction pattern for a potential interaction pattern. In addition to preventing the interface from omitting words from the search query, this approach would also help to train users from accidentally entering words. If users come to understand that the system is recording everything until they execute the search, they will also be less likely to continue inputting non-relevant speech (eg: jumping right back into a conversation or thinking out loud).

Step 1: User presses microphone icon
Step 2: Device chimes when ready
Step 3: User inputs search term
Step 4: Device chimes when it detects paus (different chime than step 2) and prompts user to confirm
Step 5: User says "search" or presses button on interface

Figure 15. Potential voice search interaction pattern

A second feature that could help prevent seniors from experiencing timeout errors is a control for setting the timeout duration. This feature could be modeled on the controls that are already widely used for setting mouse sensitivity. See Figure **16** and Figure 17. Potential timeout duration control. Rather than strictly relying on the interface to adapt to a user's cadence, this would give users control in personalizing their interface. This feature could be particularly helpful for users whose speech undergoes a dramatic change or sever speech impairment. For example, a senior with impaired speech and fine motor skills after suffering a stroke may be empowered by setting a very long timeout duration.

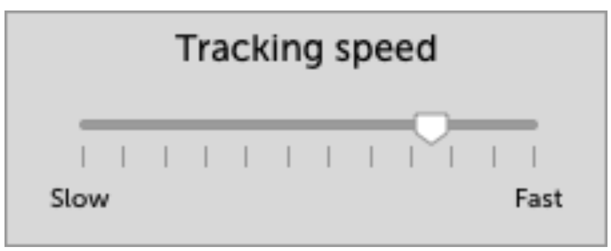

Figure 16. Example of mouse control interface

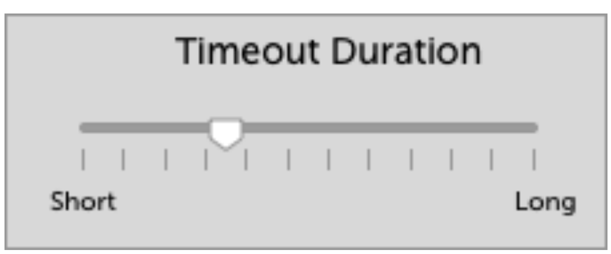

Figure 17. Potential timeout duration control

Audio Playback

If a voice search delivers search results that include a Google knowledge graph (a short summary provided by google that attempts to summarize information, answer questions and display relationships between keywords – see

Figure *18*. Example of a Google knowledge graph), the system will generate audio output to play the summary aloud (Singhal, 2017). Some seniors were strongly averse to the audio playback of results or they did not understand the purpose of the audio playback. For example, multiple seniors turned off the audio even when it was answering a question that they had searched for and asked. Another user was adamant that the audio playback was a paid advertisement. When pressed to explain why she felt this way, she compared the audio playback to pop-up video advertisements that start as soon as a user

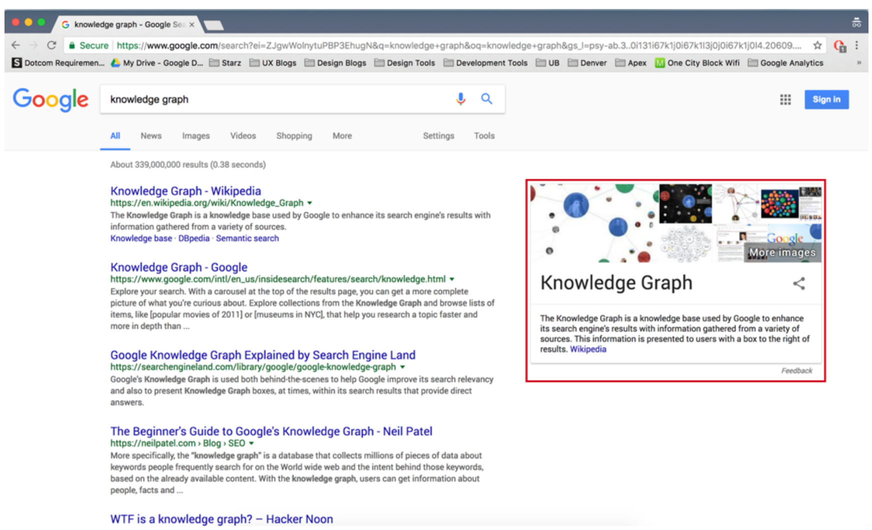

Figure 18. Example of a Google knowledge graph

visits a web page. To further complicate this issue, the setting for controlling the audio playback option is not referenced or easily accessible from the search screen. To better support seniors using voice search systems, the interface may benefit greatly from adding prompts to control non-essential playback. For example, muting audio playback triggered a prompt that asked if the user wanted to change the setting for the audio playback of results, this would promote the understanding of why the audio is playing and the user's ability to optimize their experience. Furthermore, if seniors learn to embrace the efficiencies of audio playback for search results (as they inherently did for input) this may help them avoid being distracted while seeking text based answers.

Integrating Voice Input with Visual Cues

One of the key differences between the seniors' behavior when using the KMIs compared to the VTIs was their reliance on suggested search terms. When entering text into a Google search field, the interface automatically displays suggestions to complete the query or related searches. See example in

Figure *4.* Google search engine results page with suggested search queries on computer with keyboard/mouse interface (above).

Unfortunately, once a user initiates a voice search by clicking the microphone icon, the full screen is taken over by a voice input modal. This modal is completely dedicated to the voice input and does not display suggested searches. To make things worse, the full screen take-over prevents seniors from referencing their previous search term, which increases the load on their working memory by simultaneously holding both their previous and newly desired search term.

Integrating the voice search with the traditional search interface stands to pose three key benefits to seniors: 1) seniors would be able to rely on the visual cues and free up working memory for thinking about their searches: 2) seniors would be able to use the suggested searches to form their queries, and 3) by giving the seniors useful visual information, they may be less likely to look up and away from their interfaces (which could in turn reduce the time out errors discussed above).

Current Contributions and Future Work

While the findings of this study provide a deeper understanding of how seniors interact with VUIs and the factors that shape their perceptions of those systems, the significance of the findings is restricted by the limited sample size (both in terms of numbers and diversity). Due to the functional constraints or recruiting and interviewing, the investigators were forced to rely on snowball sample of just fifteen participants. Furthermore, this sample disproportionately included affluent, white, educated participants from similar social circles. Considering that affluence and education are known to have a particularly strong effect on the adoption of technology, it is likely that the opinions of this group are not fully representative of the wider population of seniors.

Despite these limitations, the findings of this study connect strongly to pre-existing work in the field and likely address significant usability issues (which can often be identified using a sample size as low as five) (Nielsen, 2000). Future work is required to test the findings study (that seniors perceive VUIs differently in terms of learnability, usability, ease of understanding and helpfulness; and that their perceptions are shaped by keyboard familiarity, voice usability, keyboard habit, aversion to typing, voice efficiency, keyboard trust and voice novelty) and provide further insight into how these findings can best be applied to support seniors. This work could include: 1) a larger study of both seniors and younger participants; 2) the effect of more engaging research topics; 3) a deeper comparison of seniors' preferences and perceptions of VUs based on their levels of education, cognition and/or efficacy; and 4) the usability of the suggestions for designers included in this paper.

Chapter 6: Conclusion

As seniors live longer and technology becomes an even more essential component of daily life, it is imperative that designers are able to develop accessible solutions for a variety of users. Though age related changes have been well documented (particularly by Fisk (2009), there is much to learn about how seniors respond to emerging technologies. Of these emerging technologies, VUIs may be uniquely suited to help older users accommodate for age related changes to sensation, memory, attention, and movement. Furthermore, by designing interfaces that accommodate the challenges faced by seniors, they can be sure to better serve users along the entire spectrum of ability.

The goal of this study is to explore how seniors perceive VUIs and the factors that shape those perceptions. To this end, 15 seniors were asked to search for information using a traditional keyboard/mouse interface and an experimental voice/touch interface, and discuss their experiences with an investigator.

An analysis of the seniors' interview responses indicate that seniors perceive VTIs differently than traditional KMIs in terms of learnability, usability, ease of understanding and helpfulness. The VTI was also perceived as more learnable and usable, although the KMI was perceived as easier to understand, more helpful and preferable overall.

A content analysis shows that seniors' perceptions of the two interfaces were influenced by keyboard familiarity, voice usability, keyboard habit, voice usability, aversion to typing, efficiency of voice input, keyboard trust, voce novelty, keyboard speed, keyboard comfort level and fun factor (in that order).

A comparison of these results with the work of Begany et al's (2016) shows that seniors were more strongly influenced by keyboard habit, aversion to typing, and efficiency of voice input than younger users. This comparison also showed that seniors were less strongly influenced by keyboard usability, keyboard trust, voice novelty, keyboard speed, keyboard comfort level, and voice fun factor.

Observations of seniors while conducting this experiment also showed that seniors rely heavily on visual cues and were prone to distractions and time out errors. To accommodate for these behaviors, it is suggested that designers include visual cues during voice input and controls for executing voice commands and/or audio playback. Although these findings help to fill the research gap surrounding seniors and voice technology, there is still much to be learned about this topic. Our hope is that future research will be conducted to further confirm and/or refine these findings.

References

Anderson, M., & Perrin, A. (2017a, May 17). 1. Technology use among seniors. Retrieved August 19, 2017, from http://www.pewinternet.org/2017/05/17/technology-use-among-seniors/

Anderson, M., & Perrin, A. (2017b, May 17). Tech Adoption Climbs Among Older Adults. Retrieved August 19, 2017, from http://www.pewinternet.org/2017/05/17/tech-adoption-climbs-among-older-adults/

Arch, A. (2008, May 14). Web Accessibility for Older Users: A Literature Review. Retrieved from https://www.w3.org/TR/wai-age-literature/

Begany, G. M., Sa, N., & Yuan, X. (2016). Factors Affecting User Perception of a Spoken Language vs. Textual Search Interface: A Content Analysis. *Interacting with Computers*, *28*(2), 170–180. https://doi.org/10.1093/iwc/iwv029

Bernsen, N. O., & Dybkj\a er, L. (2004). Evaluation of Spoken Multimodal Conversation. In *Proceedings of the 6th International Conference on Multimodal Interfaces* (pp. 38–45). New York, NY, USA: ACM. https://doi.org/10.1145/1027933.1027941

Bourland, P. (2003). The IIR evaluation model: a framework for evaluation of interactive information retrieval systems. *Information Research*, *8*(3). Retrieved from http://www.informationr.net/ir/8-3/paper152.html

Bureau, U. C. (2014). 2014 National Population Projections Tables. Retrieved December 6, 2017, from https://www.census.gov/data/tables/2014/demo/popproj/2014-summary-tables.html

Bureau, U. C. (2017, June 22). The Nation's Older Population Is Still Growing, Census Bureau Reports. Retrieved December 6, 2017, from https://www.census.gov/newsroom/press-releases/2017/cb17-100.html

Charness, N., & Root, W. R. (2009). Aging and Information Technology Use. Retrieved August 19, 2017, from https://www.researchgate.net/publication/228633458_Aging_and_Information_Technology_Use

Chen, K., & Chan, A. (2011). A review of technology acceptance by older adults. *Gerontechnology*, *10*. https://doi.org/10.4017/gt.2011.10.01.006.00

Chisnell, D., & Redish, J. (2004, December 14). Designing Web Sites for Older Adults: A Review of Recent Research. Prepared for AARP. Retrieved from https://assets.aarp.org/www.aarp.org_/articles/research/oww/AARP-LitReview2004.pdf

Chute, D. L., & Bliss, M. E. (1994). ProsthesisWare: concepts and caveats for microcomputer-based aids to everyday living. *Experimental Aging Research*, *20*(3), 229–238. https://doi.org/10.1080/03610739408253969

Cohen, M. H., Giangola, J. P., & Balogh, J. (2004). *Voice User Interface Design*. Redwood City, CA, USA: Addison Wesley Longman Publishing Co., Inc.

Computer and Internet Use At Work Summary. (2005, August 2). Retrieved September 12, 2017, from https://www.bls.gov/news.release/ciuaw.nr0.htm

Corbett, E., & Weber, A. (2016). What Can I Say?: Addressing User Experience Challenges of a Mobile Voice User Interface for Accessibility. In *Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services* (pp. 72–82). New York, NY, USA: ACM. https://doi.org/10.1145/2935334.2935386

Czaja, S., & Chin Lee, C. (2009). Information Technology and Older Adults. In *Human-Computer Interaction* (Vols. 1–0, pp. 17–32). CRC Press. https://doi.org/10.1201/9781420088885.ch2

Czaja, S. J. (2005). The Impact of Aging on Access to Technology. *SIGACCESS Access. Comput.*, (83), 7–11. https://doi.org/10.1145/1102187.1102189

Czaja, S. J., Charness, N., Fisk, A. D., Hertzog, C., Nair, S. N., Rogers, W. A., & Sharit, J. (2006). Factors predicting the use of technology: findings from the Center for Research and Education on Aging and Technology Enhancement (CREATE). *Psychology and Aging*, *21*(2), 333–352. https://doi.org/10.1037/0882-7974.21.2.333

Davis, F. D., Bagozzi, R. P., & Warshaw, P. R. (1989). User Acceptance of Computer Technology: A Comparison of Two Theoretical Models. *Manage. Sci.*, *35*(8), 982–1003. https://doi.org/10.1287/mnsc.35.8.982

Desai, M., Pratt, L., Lentzner, H., & Robinson, K. (2001, March). Trends in Vision and Hearing Among Older Americans. CENTERS FOR DISEASE CONTROL AND PREVENTION National Center for Health Statistics. Retrieved from https://www.cdc.gov/nchs/data/ahcd/agingtrends/02vision.pdf

Ellis, L. B., Joo, H. Y., & Gross, C. R. (1991). Use of a computer-based health risk appraisal by older adults. *The Journal of Family Practice*, *33*(4), 390–394.

Employment status of the civilian noninstitutional population by age, sex, and race. (2017, February 8). Retrieved September 12, 2017, from https://www.bls.gov/cps/cpsaat03.htm

Fisk, A. D., Rogers, W. A., Charness, N., Czaja, S. J., & Sharit, J. (2009, March 26). Designing for Older Adults: Principles and Creative Human Factors Approaches, Second Edition. Retrieved August 29, 2017, from https://www.crcpress.com/Designing-for-Older-Adults-Principles-and-Creative-Human-Factors-Approaches/Fisk-Rogers-Charness-Czaja-Sharit/p/book/9781420080551

Forbes-Riley, K., & Litman, D. (2011). Designing and Evaluating a Wizarded Uncertainty-adaptive Spoken Dialogue Tutoring System. *Comput. Speech Lang.*, *25*(1), 105–126. https://doi.org/10.1016/j.csl.2009.12.002

General Demographic Questions for Any Survey. (n.d.). Retrieved July 3, 2017, from https://www.surveymonkey.com/r/SCDCJMH

Gregor, P., Newell, A. F., & Zajicek, M. (2002). Designing for Dynamic Diversity: Interfaces for Older People. In *Proceedings of the Fifth International ACM Conference on Assistive Technologies* (pp. 151–156). New York, NY, USA: ACM. https://doi.org/10.1145/638249.638277

Health, United States, 2016 - Individual Charts and Tables: Spreadsheet, PDF, and PowerPoint files. (2017, July 20). Retrieved December 6, 2017, from https://www.cdc.gov/nchs/hus/contents2016.htm

Hone, K. S., & Graham, R. (2000). Towards a Tool for the Subjective Assessment of Speech System Interfaces (SASSI). *Nat. Lang. Eng.*, *6*(3–4), 287–303. https://doi.org/10.1017/S1351324900002497

Horton, S., & Quesenbery, W. (2014). *A Web for Everyone*. Retrieved from http://rosenfeldmedia.com/books/a-web-for-everyone/

How to Write Better Demographic Survey Questions (With Examples). (n.d.). Retrieved July 3, 2017, from https://www.surveygizmo.com/survey-blog/how-to-write-better-demographic-questions/

J. Cody, M., Dunn, D., Hoppin, S., & Wendt, P. (1999). *Silver surfers: Training and evaluating Internet use among older adult learners* (Vol. 48). https://doi.org/10.1080/03634529909379178

Kamm, C., Walker, M., & Litman, D. (n.d.). *Evaluating Spoken Language Systems*.

Karsenty, L. (2002). Shifting the Design Philosophy of Spoken Natural Language Dialogue: From Invisible to Transparent Systems. *International Journal of Speech Technology*, *5*(2), 147–157.

Kim, J. (2006). Task as a Predictable Indicator for Information Seeking Behavior on the Web.

Larsen, L. B. (2003). Assessment of spoken dialogue system usability - what are we really measuring? Retrieved July 27, 2017, from

Larsen, L., & He, W. (2014, December). Older Americans With a Disability: 2008−2012. American Community Survey Reports. Retrieved from https://www.census.gov/content/dam/Census/library/publications/2014/acs/acs-29.pdf

Lee, C., Mehler, B., Reimer, B., & Coughlin, J. F. (2015). User Perceptions Toward In-Vehicle Technologies: Relationships to Age, Health, Preconceptions, and Hands-

On Experience. *International Journal of Human-Computer Interaction*, *31*(10), 667–681. https://doi.org/10.1080/10447318.2015.1070545

Leung, R., Tang, C., Haddad, S., Mcgrenere, J., Graf, P., & Ingriany, V. (2012). How Older Adults Learn to Use Mobile Devices: Survey and Field Investigations. *ACM Trans. Access. Comput.*, *4*(3), 11:1–11:33. https://doi.org/10.1145/2399193.2399195

McCreadie, C., & Tinker, A. (2005). The acceptability of assistive technology to older people. *Ageing and Society*, *25*(1), 91–110. https://doi.org/10.1017/S0144686X0400248X

Morrell, R. W. (2005). http://www.nihseniorhealth.gov: the process of construction and revision in the development of a model web site for use by older adults. *Universal Access in the Information Society*, *4*(1), 24–38. https://doi.org/10.1007/s10209-003-0085-3

Nielsen, J. (2000, March 19). Why You Only Need to Test with 5 Users. Retrieved December 2, 2017, from https://www.nngroup.com/articles/why-you-only-need-to-test-with-5-users/

Office, U. C. B. P. I. (n.d.). Nearly 1 in 5 People Have a Disability in the U.S., Census Bureau Reports - Miscellaneous - Newsroom - U.S. Census Bureau. Retrieved August 29, 2017, from https://www.census.gov/newsroom/releases/archives/miscellaneous/cb12-134.html

Pearl, C. (2016). *Designing Voice User Interfaces*. O'Reilly Media.

Plude, D. J., & Schwartz, L. K. (1996). COMPACT DISC-INTERACTIVE MEMORY TRAINING WITH THE ELDERLY. *Educational Gerontology*, *22*(6), 507–521. https://doi.org/10.1080/0360127960220601

Rogers, E. M. (2003). *Diffusion of Innovations, 5th Edition* (5th edition). New York: Free Press.

Roy, N., Stemple, J., Merrill, R. M., & Thomas, L. (2007). Epidemiology of voice disorders in the elderly: preliminary findings. *The Laryngoscope*, *117*(4), 628–633. https://doi.org/10.1097/MLG.0b013e3180306da1

Ruan, S., Wobbrock, J. O., Liou, K., Ng, A., & Landay, J. (2016). Speech Is 3x Faster than Typing for English and Mandarin Text Entry on Mobile Devices. *ArXiv:1608.07323 [Cs]*. Retrieved from http://arxiv.org/abs/1608.07323

Rubin, J., & Chisnell, D. (2008). *Handbook of Usability Testing* (Second). Wiley.

Singhal, A. (2017, December 11). Introducing the Knowledge Graph: things, not strings. Retrieved December 12, 2017, from https://googleblog.blogspot.com/2012/05/introducing-knowledge-graph-things-not.html

Smith, K. (2013). *Digital Outcasts: Moving Technology Forward without Leaving People Behind* (1 edition). Waltham, MA: Morgan Kaufmann.

Street, 1615 L., NW, Washington, S. 800, & Inquiries, D. 20036 U. 419 4300 | M. 419 4349 | F. 419 4372 | M. (2010, December 16). Generations 2010: What different generations do online. Retrieved August 19, 2017, from http://www.pewinternet.org/2010/12/16/generations-2010-what-different-generations-do-online/

Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User Acceptance of Information Technology: Toward a Unified View. *MIS Q.*, *27*(3), 425–478.

Walker, M., Litman, D., Kamm, C., & Abella, A. (1998). Evaluating spoken dialogue agents with PARADISE: Two case studies. *Computer Speech & Language*, *12*(4), 317–347. https://doi.org/10.1006/csla.1998.0110

Walters, N. (2017). *Maintaining Privacy and Security while Connected to the Internet*. AARP Public Policy Institute. Retrieved from http://www.aarp.org/content/dam/aarp/ppi/2017/08/maintaining-privacy-and-security-while-connected-to-the-internet.pdf

WHO | Life expectancy. (n.d.). Retrieved August 29, 2017, from http://www.who.int/gho/mortality_burden_disease/life_tables/situation_trends_text/en/

Wildemuth, B. M., & Freund, L. (2012). Assigning Search Tasks Designed to Elicit Exploratory Search Behaviors. In *Proceedings of the Symposium on Human-Computer Interaction and Information Retrieval* (p. 4:1–4:10). New York, NY, USA: ACM. https://doi.org/10.1145/2391224.2391228

Yankelovich, N. (1996). How Do Users Know What to Say? *Interactions*, *3*(6), 32–43. https://doi.org/10.1145/242485.242500

Appendix A: Informed Consent Form

Whom to Contact about this study:

Principal Investigator: Randall Ziman

Department: Interactive Design and Information Architecture

Telephone number: [redacted]

CONSENT FORM FOR PARTICIPATION IN RESEARCH ACTIVITIES

I. INTRODUCTION/PURPOSE: I am being asked to participate in a research study. The purpose of this study is to compare use of keyboards to use of voice/touch interfaces when conducting search tasks in users over the age of 65. I am being asked to volunteer because I am over the age of 65. My involvement in this study will begin when I agree to participate and will continue until the end of the exit interview, about 1 hour. About 15 persons will be invited to participate.

II. PROCEDURES: As a participant in this study, I will be asked to complete an entry questionnaire, receive a brief training on the keyboard and voice controls, complete a variety of search tasks (including a pre-task questionnaire for each task), complete a post system questionnaire after using each interface and complete an exit interview comparing both systems. I will be asked to come to the (location). My participation in this study will last for one visit of about an hour during which time detailed audio/video recording and note taking will occur.

III. RISKS AND BENEFITS: My participation in this study does not involve any significant risks and I have been informed that my participation in this research will not benefit me personally, but his research may have implications for the design of future spoken language search interfaces and potential improvements in the user experience of such interfaces or systems.

IV. CONFIDENTIALITY: Any information learned and collected from this study in which I might be identified will remain confidential and will be disclosed ONLY if I give permission (below). All information collected in this study will be stored in a locked home office, and electronic databases will be stored on a password protected computer. Only the investigator and members of the research team will have access to these records. If information learned from this study is published, I will not be identified by name. By signing this form, however, I allow the research study investigator to make my records available to the University of Baltimore Institutional Review Board (IRB) and regulatory agencies as required to do so by law.

Consenting to participate in this research also indicates my agreement that all information collected from me individually may be used by current and future researchers in such a fashion that my personal identity will be protected. Such use will include sharing anonymous information with other researchers for checking the accuracy of study findings and for future approved research that has the potential for improving human knowledge.

Although your confidentiality in this study is protected, confidentiality may not be absolute or perfect. There are some circumstances where research staff might be required by law to share information I have provided. For example, if an interviewer has reason to believe an elderly person is being abused (or has been abused), the interviewer is required by Maryland state law to file a report with the appropriate agencies. Similarly, if I report that I have been abused in the past, the interviewer may also have to file a report. In addition, if I am threatening serious harm to myself or another person, it may be necessary for the interviewer to warn an intended victim, notify the police or take the steps to seek hospital based treatment.

Image

☐ Yes, I give permission to use my image in scientific publications or presentations.

☐ No, I do not give permission to use my image in scientific publications or presentations

Audio

☐ Yes, I give permission to use my voice in scientific publications or presentations.

☐ No, I do not give permission to use my voice in scientific publications or presentations

V. **SPONSOR OF THE RESEARCH:** This research study is for a master's thesis.

VI. **COMPENSATION/COSTS:** My participation in this study will involve no cost to me.

VII. **CONTACTS AND QUESTIONS:** The principal investigator(s), Randall Ziman (student researcher) and Greg Walsh (faculty advisor) has offered to and has answered any and all questions regarding my participation in this research study. If I have any further questions, I can contact Randall Ziman (student researcher) at [redacted] or [redacted] or Greg Walsh (faculty advisor) at 410 837 5473.

For questions about rights as a participant in this research study, contact the UB IRB Coordinator: 410-837-6199, irb@ubalt.edu.

VIII. **VOLUNTARY PARTICIPATION** I have been informed that my participation in this research study is voluntary and that I am free to withdraw or discontinue participation at any time.

I will be given a copy of this consent form to keep.

IX. **SIGNATURE FOR CONSENT** The above-named investigator has answered my questions and I agree to be a research participant in this study. By signing this consent form, I am acknowledging that I am at least 18 years of age.

Participant's Name: ________________________________ Date: _______________

Participant's Signature: _____________________________ Date: _______________

Investigator's Signature: _____________________________ Date: _______________

Appendix B: Background Questionnaire

Name: ________________________

What is your gender?

- ❏ Female
- ❏ Male

What is your age?

- ❏ 65-69
- ❏ 70-74
- ❏ 75-79
- ❏ 80-85
- ❏ 86 or older

What is the highest level of education you have completed?

- ❏ High school or equivalent
- ❏ Vocational/technical school (2 year)
- ❏ Some college
- ❏ Bachelor's degree
- ❏ Master's degree
- ❏ Doctoral/Professional degree (MD, JD, etc.)
- ❏ Other

What is your annual salary (or what was your salary at the time you retired)?

- ❏ Under $20,000
- ❏ $20,000 - $34,999
- ❏ $35,000-$49,999
- ❏ $50,000 - $74,999
- ❏ $75,000 - $99,999
- ❏ Over 100,000
- ❏ Would rather not say

Ethnicity: We want to make sure that we have spoken to a broad mix of people in your area. Are you, yourself, of Hispanic origin or descent, such as Mexican, Puerto Rican, Cuban or other Spanish background?

- ❏ Yes
- ❏ No

Race: How would you describe yourself?

- ❏ American Indian or Alaska Native
- ❏ Asian
- ❏ Black or African American
- ❏ Native Hawaiian or Other Pacific Islander
- ❏ White
- ❏ Other: ___________________

How would you rate your ability with using a computer?

- ❏ Novice
- ❏ Intermediate
- ❏ Advanced

How many hours a day do you use a computer?

- ❏ Under 1
- ❏ 2-4
- ❏ 5-10
- ❏ 11 or more

How would you rate your ability with using a tablet or smartphone?

- ❏ Novice
- ❏ Intermediate
- ❏ Advanced

How many hours a day do you use a tablet or smartphone?

- ❏ Under 1
- ❏ 2-4
- ❏ 5-10

- ❏ 11 or more

What activities do you do online (check all that apply)?

- ❏ Connecting with friends and family
- ❏ Email
- ❏ Shopping
- ❏ News or learning
- ❏ Other: __________________

How often do you search for information online?

- ❏ Never
- ❏ A few times a month
- ❏ A few times a week
- ❏ Every day
- ❏ Multiple times a day

What devices do you usually use when you are looking for information online?

- ❏ Computer
- ❏ Tablet
- ❏ Smart Phone
- ❏ Other
- ❏ Not Applicable

What types of information do you usually search for (check all that apply)?

- ❏ News
- ❏ Finance
- ❏ Sports
- ❏ Politics
- ❏ Entertainment
- ❏ Factual/Reference
- ❏ Other: _______________
- ❏ Not Applicable

Have you ever used a touch screen?

- ❏ Yes
- ❏ No

If you have used a touch screen, how long have you been using it?

- ❏ Less than 6 months
- ❏ 6 to 12 months
- ❏ 1 to 3 years
- ❏ 3+years
- ❏ Not sure
- ❏ Not Applicable

How often do you use touch screens?

- ❏ Never
- ❏ A few times a month
- ❏ A few times a week
- ❏ Every day
- ❏ Multiple times a day

Have you ever used voice activated technology (Siri, voice-to-text, Amazon Echo, Google Home, etc.)?

- ❏ Yes
- ❏ No

If you have used voice activated technology, how long have you been using it?

- ❏ Less than 6 months
- ❏ 6 to 12 months
- ❏ 1 to 3 years
- ❏ 3+years
- ❏ Not sure
- ❏ Not Applicable

How often do you use voice activated technology?

- ❏ Never
- ❏ A few times a month

- ❏ A few times a week
- ❏ Every day
- ❏ Multiple times a day

Have you ever used the voice search function of a search engine?

- ❏ Yes
- ❏ No

If you have used the voice search function, how long have you been using it?

- ❏ Less than 6 months
- ❏ 6 to 12 months
- ❏ 1 to 3 years
- ❏ 3+years
- ❏ Not sure
- ❏ Not Applicable

How often do you use the voice search function?

- ❏ Never
- ❏ A few times a month
- ❏ A few times a week
- ❏ Every day
- ❏ Multiple times a day

Appendix C: Search Tasks

Factual Tasks

FT1 Movies

- Topic: Recently, a friend told you about a movie that he liked. It starred a French actress. The title was a woman's name. It was an award-winning film that came out in 2001. This actress also played the lead role in a movie with Tom Hanks in 2006.
- Task: Find the name of the actress and the name of both movies. Save the document(s) where you have found the information required.

FT2 Cruise Ships

- Topic: Recently, you were talking with a friend about going on a cruise. She has been raving for years about the last cruise she went on and you would like to book with the same company. She doesn't remember the name of the company, but she does remember that in 2006, the cruise ship was the largest in the world.
- Task: Find the cruise company that had the largest cruise ship in 2006. Save the document(s) where you have found the information required.

FT3 Elections

- Topic: Recently, you were debating about which state has the highest voter turnout in the last election. A friend of yours said it was Minnesota, but you thought that it was somewhere on the east coast.
- Task: Find out which state had the highest voter turnout for the 2016 presidential election. Save the document(s) where you have found the information required.

FT4 Valuable Metals

- Topic: Recently, you have been hearing a lot about stock market alternatives. You are considering investing in platinum but you want to learn more about its historical performance
- Task: Find out how much the value of an ounce of platinum has increased or decreased in the last 3 years. Save the document(s) where you have found the information required.

Interpretive Tasks

IT1 Organic Food

- Topic: A friend of yours insists that you must only buy and eat organic foods. She has been warning you about genetically modified foods and their harmful effects. You have also heard of people who only eat raw foods in their diet. You have decided you need to find some information on organic food, genetically modified food and raw food to be able to discuss this further.
- Task: Find the benefits and/or harmful effects of each type of food. Save the document(s) where you have found the information required.

IT2 Solar energy

- Topic: You recently watched an interview about the decline of the US coal industry with some friends. One of your friends insisted that solar energy will be wave of the future. Another friend was very skeptical and thought that the costs outweighed the benefits.
- Task: Find out the pros and cons of solar energy. Save the documents where you have found the information required.

IT3 K12 Education

- Topic: Recently, a friend of yours was talking about his finding a new school for his Grandson. Their family is trying to decide between sending him to a charter school, a public school or a private school. You have decided that you would like to learn more about the differences between each type of school.

- Task: Find out about the advantages and disadvantages of each type of school (charter, public, private). Save the document(s) where you have found the information required.

IT4 High School Extracurricular Activities

- Topic: Recently, you heard that a local school district is thinking about cutting their sports programs because of lack of funding. Some school board members are saying that the cuts are a good thing because they think it will remove distractions. However, the principal is concerned that this will have a negative effect on the local high school.
- Task: Find the benefits and/or harmful effects high school sports programs. Save the document(s) where you have found the information required.

Exploratory Task

ET1 Vegetarianism

- Topic: You have friends who are vegetarians and some who are vegans. You want to have a better understanding why your friends have chosen to not eat meat and/or animal products.
- Task: You want to understand why some people eat meat and some do not. You want to know if it is a cultural thing or something else. Save the document(s) where you have found the information required.

ET2 Exercise plan

- Topic: A friend of yours who just turned 90 was told by her doctor that getting more low-impact exercise will increase her fitness and help her avoid injuries. Your friend does not use the Internet and has asked you to help create an exercise program for her.

- Task: Put together two thirty-minute low-impact exercise programs that she could alternate between during the week. Save the document(s) where you have found the information required.

ET3 Solar heating

- Topic: Some of your family members are planning to build a new house and have heard that using solar energy panels for heating can save a lot of money. Since they do not know anything about home heating and the issues involved, they have asked for your help.
- Task: Do some research to identify some issues that need to be considered in deciding between more conventional methods of home heating and solar panels. Save the document(s) where you have found the information required.

ET4 Racial Profiling

- Topic: You are taking a course on modern issues in America. You've been given an assignment on racial profiling, and are expected to write a paper on it. You decide to begin by trying to understand what racial profiling is, and explore and examine the issues, organizations and laws concerning it.
- Task: Use the internet to find two possible topics for your paper. Then find three sources that you could use to learn more about each topic.

Appendix D: Post System Interview

How easy/difficult did you find learning to use [the system]?

- ❑ 1 – Very difficult
- ❑ 2 – Difficult
- ❑ 3 – Neutral
- ❑ 4 – Easy
- ❑ 5 – Very easy

How easy/difficult did you find using [the system] to search for information?

- ❑ 1 – Very difficult
- ❑ 2 – Difficult
- ❑ 3 – Neutral
- ❑ 4 – Easy
- ❑ 5 – Very easy

How easy/difficult did you find understanding [the system] interface?

- ❑ 1 – Very difficult
- ❑ 2 – Difficult
- ❑ 3 – Neutral
- ❑ 4 – Easy
- ❑ 5 – Very easy

How useful did you find [the system]?

- ❑ 1 – Not at all useful
- ❑ 2 – Slightly useful
- ❑ 3 – Somewhat useful
- ❑ 4 – Very useful
- ❑ 5 – Extremely useful

Appendix E: Exit Interview

1. How would you compare the overall differences between the keyboard and the voice search?
2. Which system did you find more helpful? Why?
3. Which system was easier to learn to use? Why?
4. Which system was easier to actually use? Why?
5. Which system made it more difficult to complete your tasks? Why?
6. Which system gave you search results that you were more satisfied with? Why?
7. Which system conducted better searches? Why?
8. Which system did you like best overall? Why?
9. What features of completing these tasks did you like the most? Why?
10. What features of completing these tasks did you dislike the most? Why?
11. What features do you think could be added to these systems in the future? Why?

Appendix F: Example Interview Script for Keyboard First Protocol

Introduction

Hello ____________, thank you for coming. My name is _________ and I'm very glad that you have agreed to take part in our research study about technology. During this session, I will be working from a script to ensure that my instructions to everyone who participates in the study are the same. Also, I hope you don't mind, but I'm going to use my computer to take notes because it helps me with analyzing the session.

Before we begin, I would like to ask you to sign this consent form. This form explains the purpose and procedure of our study and asks for your permission to record video and audio from our session.

[User signs consent form – see attachments]

Thank you much. Now, while I get our recording equipment turned on, can you please fill out this background questionnaire? Your answers here will help us look for trends in people's responses

[While researcher begins recording, ask user to complete entry questionnaire – see attachments]

[Researcher begins recording audio using phone]

Next, I have some information for you:

I am here to learn about how people use voice controlled devices and how that is different than using keyboard controlled devices. During this session, I will be asking you to search for different types of information and then observing you as you complete the task.

While you complete the tasks, please try to think aloud as much as possible (but you don't need to worry about that if you are using your voice to conduct the searches). Just tell me whatever is going through your mind. Please know that we're not testing you, and there is no such thing as a wrong answer. Your responses will help us understand what works or doesn't work about the technology.

Also, I would like you to know that I am an independent researcher who had nothing to do with the design of the technology you're about to try out, so please be honest in your feedback.

Today, you will be assigned 6 tasks that require you to search for information online and bookmark valuable pages. For first 3 tasks, you will be asked to use a computer with a traditional keyboard and mouse. For the other 3 tasks, you will be asked to use the voice search on an iPad. For each task, you will be given a situation and then asked to search for information about that subject. When you find information that will be helpful, you will bookmark the page.

The whole session will take about 1 hour.

Do you have any questions before we begin?

[answer questions]

If you have any doubts during the test, feel free to ask me any questions. However, please remember that I might not be able to answer them during the test. We want to learn about your experience with the site, so I can't guide you. I'll be glad to answer all your questions after the test.

System A - Training

You are probably familiar with this device, but to ensure that all participants start with similar knowledge about the technology we will be using, I will conduct a brief orientation.

[Researcher will demonstrate capabilities by searching for 'computers']

We will begin with the keyboard. To conduct a keyboard search with the computer:

- Use the mouse to place your cursor over the search box and click once
- Use the keyboard to enter your search
- The, to run the search, hit enter on the keyboard or use the mouse to click on the 'google search button'
- When your search returns results, you can click on each item to visit the actual page

- You can modify your search by changing the search terms in the search box at the top of the page and clicking enter/the google search button
- You can move forward and back using the navigation arrows in the upper left-hand corner
- You can start over by returning to google.com or clicking the google icon

When you find a page that you think is helpful bookmark the page by

- clicking the star icon in the url bar and then clicking the blue done button
- Click the blue done button

Now, let’s have you do a practice search task.

- Topic: Recently, you were debating about which state has the highest voter turnout in the last election. A friend of yours said it was Minnesota, but you thought that it was somewhere on the east coast.
- Task: Find out which state had the highest voter turnout for the 2016 presidential election. Save the document(s) where you have found the information required.

[Participants conducts practice search]

Now that you have done a practice search, let’s move on.

System A - Tasks

Task 1

Before you start your search task, we would like to learn more about your previous thoughts and experience coming into the activity

How familiar are you with the topic of movies?

1. Not at all familiar
2. Slightly familiar
3. Somewhat familiar
4. Very familiar
5. Extremely familiar

How interesting do you find the topic of movies?

1. Not at all interesting
2. Slightly interesting

3. Somewhat interesting
4. Very interesting
5. Extremely interesting

How easy/difficult do you think it will be to find information about this topic?

1. Very difficult
2. Difficult
3. Neutral
4. Easy
5. Very easy

Now that we have established a baseline for your thoughts and experiences about this topic, let's move on to the actual search.

- Topic: Recently, a friend told you about a movie that he liked. It starred a French actress. The title was a woman's name. It was an award-winning film that came out in 2001. This actress also played the lead role in a movie with Tom Hanks in 2006.
- Task: Find the name of the actress and the name of both movies. Save the document(s) where you have found the information required.

Task 2

Now, let's move on to a new task.

How familiar are you with the topic of organic foods?

1. Not at all familiar
2. Slightly familiar
3. Somewhat familiar
4. Very familiar
5. Extremely familiar

How interesting do you find the topic of organic foods?

1. Not at all interesting
2. Slightly interesting
3. Somewhat interesting

4. Very interesting
5. Extremely interesting

How easy/difficult do you think it will be to find information about this topic?

1. Very difficult
2. Difficult
3. Neutral
4. Easy
5. Very easy

Now, let's move on to the actual search.

- Topic: A friend of yours insists that you must only buy and eat organic foods. She has been warning you about genetically modified foods and their harmful effects. You have also heard of people who only eat raw foods in their diet. You have decided you need to find some information on organic food, genetically modified food and raw food to be able to discuss this further.
- Task: Find the benefits and/or harmful effects of each type of food. Save the document(s) where you have found the information required.

Task 3

Now, let's move on to a new task.

How familiar are you with the topic of vegetarianism?

1. Not at all familiar
2. Slightly familiar
3. Somewhat familiar
4. Very familiar
5. Extremely familiar

How interesting do you find the topic of vegetarianism?

1. Not at all interesting
2. Slightly interesting
3. Somewhat interesting
4. Very interesting

5. Extremely interesting

How easy/difficult do you think it will be to find information about this topic?

1. Very difficult
2. Difficult
3. Neutral
4. Easy
5. Very easy

Now, let's move on to the actual search.

- Topic: You have friends who are vegetarians and some who are vegans. You want to have a better understanding why your friends have chosen to not eat meat and/or animal products.
- Task: You want to understand why some people eat meat and some do not. You want to know if it is a cultural thing or something else. Save the document(s) where you have found the information required.

System A - Post System Interview

How easy/difficult did you find learning to use the keyboard?

1. Very difficult
2. Difficult
3. Neutral
4. Easy
5. Very easy

How easy/difficult did you find using the keyboard to search for information?

1. Very difficult
2. Difficult
3. Neutral
4. Easy
5. Very easy

How easy/difficult did you find understanding the keyboard interface?

1. Very difficult

2. Difficult
3. Neutral
4. Easy
5. Very easy

How useful did you find the keyboard?

1. Not at all useful
2. Slightly useful
3. Somewhat useful
4. Very useful
5. Extremely useful

System B - Training

Now we will switch to our second device.

[Researcher will demonstrate capabilities by searching for 'computers']

To voice search with the tablet:

- Touch the microphone icon in the bar at the top of the page to initiate a voice search
- When you stop talking, the search will automatically run
- When your search returns results, you can click on each item to visit the actual page
- You can move forward and back using the navigation arrows in the upper left-hand corner
- You can start over by returning to google.com or clicking the google icon
- You can modify your search by touching the microphone in the bar at the top of a google search page (you might need to use the back arrows to navigate back to a google page)

When you find a page that you think is helpful bookmark the page by

- clicking the star icon in the upper right-hand corner

System B - Tasks

Task 4

Now, let's move on to a new task.

How familiar are you with the topic of cruise ships?

1. Not at all familiar
2. Slightly familiar
3. Somewhat familiar
4. Very familiar
5. Extremely familiar

How interesting do you find the topic of cruise ships?

1. Not at all interesting
2. Slightly interesting
3. Somewhat interesting
4. Very interesting
5. Extremely interesting

How easy/difficult do you think it will be to find information about this topic?

1. Very difficult
2. Difficult
3. Neutral
4. Easy
5. Very easy

Now, let's move on to the actual search.

- Topic: Recently, you were talking with a friend about going on a cruise. She has been raving for years about the last cruise she went on and you would like to book with the same company. She doesn't remember the name of the company, but she does remember that in 2006, the cruise ship was the largest in the world.
- Task: Find the cruise company that had the largest cruise ship in 2006. Save the document(s) where you have found the information required.

Task 5

Now, let's move on to a new task.

How familiar are you with the topic of solar energy?

6. Not at all familiar
7. Slightly familiar
8. Somewhat familiar
9. Very familiar
10. Extremely familiar

How interesting do you find the topic of solar energy?

6. Not at all interesting
7. Slightly interesting
8. Somewhat interesting
9. Very interesting
10. Extremely interesting

How easy/difficult do you think it will be to find information about this topic?

6. Very difficult
7. Difficult
8. Neutral
9. Easy
10. Very easy

Now, let's move on to the actual search.

- Topic: You recently watched an interview about the decline of the US coal industry with some friends. One of your friends insisted that solar energy will be wave of the future. Another friend was very skeptical and thought that the costs outweighed the benefits.
- Task: Find out the pros and cons of solar energy. Save the documents where you have found the information required.

Task 6

Now, let's move on to a new task.

How familiar are you with the topic of fitness?

1. Not at all familiar
2. Slightly familiar

3. Somewhat familiar
4. Very familiar
5. Extremely familiar

How interesting do you find the topic of fitness?

1. Not at all interesting
2. Slightly interesting
3. Somewhat interesting
4. Very interesting
5. Extremely interesting

How easy/difficult do you think it will be to find information about this topic?

1. Very difficult
2. Difficult
3. Neutral
4. Easy
5. Very easy

Now, let's move on to the actual search.

- Topic: A friend of yours who just turned 90 was told by her doctor that getting more low-impact exercise will increase her fitness and help her avoid injuries. Your friend does not use the Internet and has asked you to help create an exercise program for her.
- Task: Put together two thirty-minute low-impact exercise programs that she could alternate between during the week. Save the document(s) where you have found the information required.

System B - Post System Interview

How easy/difficult did you find learning to use the voice and touch system?

1. Very difficult
2. Difficult
3. Neutral
4. Easy

5. Very easy

How easy/difficult did you find using the voice and touch system to search for information?

1. Very difficult
2. Difficult
3. Neutral
4. Easy
5. Very easy

How easy/difficult did you find understanding the voice and touch system interface?

1. Very difficult
2. Difficult
3. Neutral
4. Easy
5. Very easy

How useful did you find the voice and touch system?

1. Not at all useful
2. Slightly useful
3. Somewhat useful
4. Very useful
5. Extremely useful

Exit Interview

Thank you for taking the time today to show us how you search for information. Before we conclude, we would like to ask you a few questions about your experience here today:

How would you compare the overall differences between the keyboard and the voice search?

Which system did you find more helpful? Why?

Which system was easier to learn to use? Why?

Which system was easier to actually use? Why?

Which system made it more difficult to complete your tasks? Why?

Which system gave you search results that you were more satisfied with? Why?

Which system conducted better searches? Why?

Which system did you like best overall? Why?

What features of completing these tasks did you like the most? Why?

What features of completing these tasks did you dislike the most? Why?

What features do you think could be added to these systems in the future? Why?

Conclusion

Thank you again for your time today, I can assure you that we are very excited to learn more about how people interact with technology and we hope that improving our understanding will allow us to design better products for everyone.

[Participant Leaves]