AGING IN DIGITAL WORLDS:
WHAT IS THE EXPERIENCE OF SENIORS
USING ACCESSIBLE GOVERNMENT WEBSITES?

KENNETH HOWARD ANDERSON

A DISSERTATION SUBMITTED TO
THE FACULTY OF GRADUATE STUDIES
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

GRADUATE PROGRAM IN EDUCATION
YORK UNIVERSITY
TORONTO, ONTARIO

February 2016

Abstract

Previous research has shown that seniors often face barriers to internet website accessibility due to age-related physical disabilities and memory loss issues. Web content accessibility guidelines (W3C WCAG 2.0, 2008) have been employed by government legislation to assist in making websites accessible to people with disabilities, including seniors suffering from age-related ability changes. Automated accessibility software checkers are used to test the technical accessibility of websites conforming to these guidelines. This study examined how usable seniors found these technically accessible websites, adopting a mixed methods approach to gather data on the user experience of seniors while accessing technically accessible websites. Quantitative data were gathered to provide a broad perspective on the user experience of the participants, while qualitative data were analyzed to determine themes and relationships within the user experiences of the participants. The quantitative analysis is integrated with the qualitative analysis to produce inferences, descriptions, and conclusions about the user experiences of the participants. The study provided evidence as to whether technical accessibility standards met the needs of seniors while accessing government services on the internet, while documenting the reported experiences of seniors using those websites.

Acknowledgements

I think of the PhD quest as a grand journey, beginning as it does with an inquiring mind, and reaching a station of sorts with the completion of a document such as this. The journey cannot be undertaken without the helpful involvement of others. I am most grateful for the assistance, support, and guidance provided by my supervisor Dr. Ron Owston, and my committee members Dr. Gail Mitchell and Dr. Sue Winton. I thank my external examiner Dr. Clare Brett, examination chair Dr. Theresa Shanahan, and internal member David Gelb, for their wise input and attention to my work.

I thank my family for their involvement in this journey - my parents Muriel and John, and my brothers David and Donald, for their inspiration. I thank my children Lyndsey, Bradley, and Natalie and their families for their support throughout. I would not have commenced nor completed my quest without the wise counsel of Xiaohui Gao, who accompanied me along the way and nudged me forward when I needed it most. In many ways, my journey is our journey.

Table of Contents

List of Tables

List of Figures

Chapter 1 - Introduction

This chapter introduces my central topic and its significance. The aim of the study is outlined and substantive, theoretical, and methodological objectives are discussed. Finally, key elements of the context within which the study is situated are noted.

Description and Significance of the Topic

Today's digital worlds of web-provided information, interaction, and services mean that we now age in our abilities to conjoin, interact, and access these worlds, through declines in our perceptual capacities (see Newell, 2008). For example, 14% of Canadians report having a disability, and this number increases to 43.2% for those over the age of 65 (HRSDC, 2009). The benefits of digital worlds for seniors include quality of life and independence through service and information access (Czaja, Gregor & Hanson, 2009). Exclusion from the benefits of these worlds awaits those who are unable to access them. A specific concern is the possibility of exclusion from digitally mediated service access, for those of us suffering from age-related disabilities. This research study centres on the topic of government website usability for seniors who are 55 years and older, including those that may report disabilities. It is focused on measuring the experiences of seniors as they access and perform tasks on Ontario government websites that meet current accessibility standards. The study assessed the experiences by means of participant observation, survey questionnaires and semi-structured interviews of a selected group of seniors.

Numerous guidelines exist for the design of websites that are both usable and accessible to older users (W3C, 2008). Internet websites often incorporate policies that claim support for the full inclusion of persons with disabilities, including age-related disabilities. For example, the Ontario Public Service has an accessible customer service policy created to comply with human

rights codes and accessibility legislation in Canada. The policy's goal is to make Ontario barrier-free by 2025, and ensure that all government services become progressively more accessible (Ontario Public Service, 2011). Legislation passed in the province of Ontario has led to the development of these types of policies (Accessibility for Ontarians with Disabilities Act, 2005).

It is imperative that research be conducted in an ongoing and timely fashion in order to enable the steady progression to accessibility. It is imperative that seniors not be excluded from accessing government services offered on the internet. Government regulation mandates that all new Ontario government websites must meet international accessibility standards by January 1, 2012, with other new public sector and large organizations meeting this standard by January 1, 2014. Existing websites must conform by 2016 and 2021, respectively (Ontario Regulation 191/11, 2011). Government websites providing service must be usable and accessible to seniors in order to not exclude them from using the services provided therein.

I believe this study is particularly relevant as the proportion of seniors in the population is increasing (HRSDC, 2009) and more government information and services become available online (e.g. Money, Lines, Fernando, & Elliman, 2011; Sayago & Blat, 2010). The research contributes to the literature on the design of accessible websites for seniors. It also informs government policy on whether existing standards for website accessibility are sufficient and/or relevant to meeting the needs of seniors.

Study Objectives

There were three primary goals to my study. These goals were grouped into the categories of substantive, theoretical, and methodological objectives. I employed the logical argument structure presented in Wallace and Wray (2009).

Substantive objective

The main goal directing the study was to explore and determine how usable government websites are to seniors. The government websites that were tested in the study were those that met most technical accessibility standards as defined by guidelines developed to primarily address the needs of disabled persons. These guidelines suggest that older persons will also benefit from accessible websites due to the challenges faced while aging (W3C WAI, 2005a). While joining together the concepts of accessibility and usability I aimed to emphasize the user experience of seniors accessing government websites. I explored usability and user experience within the environment of accessible government websites.

Accessibility is often seen as a sibling to usability, and/or a requisite ingredient of a website before it can even be used (Rubin & Chisnell, 2009). The examination of usability through user experience includes the need to reference accessibility. A conflicting perspective is that legislation focuses on accessibility and not usability. It was my aim to bring these two conceptual models together by examining relations between them in this study, toward an objective of determining whether accessibility compliance is sufficient to meet the needs of seniors accessing government websites.

Theoretical objective

A theoretical objective of the study was to integrate conceptual elements from the characteristics of seniors, usability research, and user experience theory. Cognitive load theory (CLT) (Sweller, Merrienboer, & Paas, 1998) can be useful in describing some characteristics of seniors, particularly as the theory relates to working memory, its limitations, and the load placed upon it by learning needs (as related to websites in this study). Linking cognitive load theory

with usability in the design of education software is suggested in the literature, for example as proposed in Hollender, Hofmann, Deneke, and Schmitz (2010). However, I did not utilize a full synthesis here. Rather, I am acknowledging that elements of cognitive load theory were insightful for informing my study, in particular in their relationship to potentially understanding the characteristics of seniors.

Usability concepts, especially those focusing on user satisfaction, are critical in determining the user experience of persons using websites. I used elements of user experience theory in my study to help inform the findings from the gathered data. User experience theory is sometimes seen as an extension of usability theory and research. This study hopes to contribute to theory by highlighting a need to employ user experience concepts when determining the usability of websites, particularly in the context of seniors and their usage.

Methodological objective

Methodological approaches included designing a mixed-methods research investigation. The design enabled the gathering of both quantitative and qualitative data. Quantitative data included the gathering of numeric accessibility data on government websites along with numeric measures of the experiences of seniors using them. Qualitative data included the gathering of narrative data from the seniors in the form of observations and interview responses. Data collected from both qualitative (QUAL) and quantitative (QUAN) instruments provided for the drawing of inferences using triangulation of the data. The quality of the inferences arising from the individual QUAL and QUAN strands were assessed using an integrative framework (Teddlie & Tashakkori, 2009). In this framework, design quality and interpretive rigour were the criteria used in assessing the inference quality of the study.

Context of the Study

Two contextual particulars are discussed in this section. The age demarcation for seniors is discussed and a reference to the heterogeneity of the group is made. The legislation on accessibility in the Province of Ontario is reviewed, along with the conformance guideline (WCAG 2.0) that is used in the legislation to ensure accessible websites.

Seniors

Seniors are defined variously in the literature as those over age 55, 60, 65, etc. (for examples see Hanson, 2011; Veenhof & Timusk, 2009). The term 'older adult' is likewise defined variously depending on the context being considered (Wagner, Hassainein, & Head, 2010). For the purposes of this study, I used the senior demarcation age found in the environment in which the study is conducted. In this study, the environment was a Seniors Centre with an age demarcation of 55 years and older. Snowball sampling was conducted from this population and is discussed in further detail in the methods section of this proposal.

Seniors can be roughly classified into three categories: fit older people, frail older people, and disabled people (Newell, 2011). Within categories such as these, seniors are often described as possessing a somewhat homogenous set of characteristics that are considered to likely impact their ability to access the internet. Age-related declines in functional abilities such as perception, movement, and cognition are often generally ascribed to seniors (e.g. Hanson, 2011). However, it is argued that seniors are more a heterogeneous group (e.g. Affonso de Lara, Watanabe, Beletato dos Santos & Fortes, 2010). In particular, many older adults do not suffer with cognitive decline (Czaja, Sharit, Hernandez, Nair & Lowenstein, 2010) and an argument can be made that stereotyping seniors' characteristics for research purposes can have a negative

impact (Dickinson, Arnott & Prior, 2007). Broad-based theories on the existence or non-existence of cognitive decline exist in the literature, but I have chosen to not directly address them in this study.

Accessibility for Ontarians with Disabilities Act, 2005

The Accessibility for Ontarians with Disabilities Act, 2005 (AODA) and website accessibility regulation (Ontario Regulation 191/11, 2011) provide a legislative framework for accessibility to internet resources for persons with disabilities in the Province of Ontario. The legislation and regulation require government institutions and other organizations to conform to accessibility guidelines within time frames assigned to the specific category of institution or organization. The Government of Ontario requires its new websites achieve conformance on January 1, 2012, and its existing websites by January 1, 2016.

The Government of Ontario, through its legislation and regulation thereto, has used the term *accessibility standard* as defining a process that is required to "…set out measures, polices, practices or other requirements for the identification and removal of barriers with respect to goods, services, facilities…" etc. (Accessibility for Ontarians with Disabilities Act, 2005). The standard chosen for the Government of Ontario (and its institutions) to meet its requirements within the act, regulation, and process is "…the World Wide Web Consortium Web Content Accessibility Guidelines (WCAG) 2.0, at Level AA…" (Ontario Regulation 191/11, 2011).

WCAG 2.0 is a conformance guideline that provides recommendations to make websites more accessible through attention to four principles. It is a recommendation of the World Wide Web Consortium (W3C). The principles require websites to be perceivable, operable, understandable, and robust (W3C WCAG 2.0, 2008). Within each principle are a number of

guidelines that provide goals for achieving the principles. Success criteria are embedded within each guideline, and provide a testable means (including levels of conformance such as the level AA chosen by the Ontario Government) of achieving conformance. Automated software tools such as AChecker (2015a) are available to conduct the testing. The web accessibility initiative of the W3C is concerned with ensuring that persons with disabilities have web access, while suggesting that older persons will benefit from accessible artefacts due to the physical and mental declines experienced in human aging (W3C WAI, 2005a). The use of these guidelines to assess the websites in my study helps to provide an indication of the characteristics of the digital environment of the study. The use also provides a measure of the success (or failure) of the websites to comply with the guidelines, given the existence of legislation in that regard.

Summary

Legislation in the Province of Ontario requires organizations and government institutions to conform to web content accessibility guidelines for both existing and new websites within specific time frames as set out in the regulation. In addition to directly addressing the needs of identified disabilities, WCAG 2.0 conformance guidelines are intended to help seniors faced with changes brought about by aging. Testing of and conformance to WCAG 2.0 is required for Ontario Government websites, towards the goal of increasing accessibility for Ontarians. Increasing accessibility for Ontarians is a government goal, in line with the economic context of the province. For example, the Ontario Ministry of Economic Development, Employment & Infrastructure argues that accessibility legislation benefits everyone in Ontario, and that the demographics are such that businesses cannot overlook the need for accessibility. It is this Ministry's finding that "In the next 20 years, an aging population and people with disabilities will represent 40% of total income in Ontario" (Ontario Ministry of Economic Development,

Employment & Infrastructure, 2015a). Further, accessibility also benefits seniors, and, "…beyond being good for business – it's just the right thing to do" (Ontario Ministry of Economic Development, Employment & Infrastructure, 2015b, p.1.). Accessibility is defined by this Ministry as "Accessibility simply means giving people of all abilities opportunities to participate fully in everyday life" (Ontario Ministry of Economic Development, Employment & Infrastructure, 2015a). The economic importance of accessibility is emphasized by noting that "With the aging population, consumers are increasingly represented by the disability community" (Ontario Ministry of Economic Development, Employment & Infrastructure, 2015b, p.1.).

My study tested websites for accessibility conformance and compared the results with the reported user experience of seniors, while they performed information-seeking tasks on those websites. The theoretical aim of the study is to employ and measure concepts from the fields of usability research and user experience theory. From these fields I intend to emphasize user experience concepts, to explore how seniors think, act, and feel about their experiences accessing and using government websites that have met accessibility conformance guidelines.

Chapter 2 - Literature Review

Introduction

This chapter reviews literature on accessibility and usability of online services for seniors. I argue that accessibility and usability measures benefit from including the affective domain – an emphasis found in literature on user experience. I take the position that user experience is critical to a determination of usability by seniors, notwithstanding that accessibility may be required before any usage of a website can be made. I provide a conceptual position that includes concepts from the field of user experience and the practice of usability study.

A first step towards adopting this conceptual position is to adopt a perspective on experience. John Dewey (1938/1997) posited that experience includes the criteria of continuity and interaction. Continuity means that an experience "…both takes up something from those which have gone before and modifies in some way the quality of those which come after" (Dewey, 1938/1997, p.35). Interaction means that an experience is an interplay between internal and external conditions, "…a transaction taking place between an individual and…his environment…" (Dewey, 1938/1997, p.43). Dewey uses these concepts in an examination of their importance to proposed educational goals. These terms can also have utility by abstracting them to understand user experience in the context of interactions with digital technologies. The transactions between a person and a digital environment can also be thought of as being influenced by the past and influencing the future.

Marc Hassenzahl (2014, p.7) considers user experience to be "…about creating a meaningful experience through a device", with experience being thought of as the story of a person's interaction with their world. The focus is less on the knowledge gained from the

experience and more on the meaningfulness of the experience. Meaningfulness, for Hassenzahl, is the situation where "…emotions and fulfilment of universal psychological needs…have an accentuated role" (Hassenzahl, 2014, p.8). This accentuated role results in a design approach that determines needs and emotions first, then works on the functionalities involved, resulting in "…products which are sensitive to the particularities of human experience…products able to tell enjoyable stories through their use or consumption" (Hassenzahl, 2014, p. 15). In Hassenzahl, Eckoldt, Diefenbach, Laschke, Lenz, and Kim (2013), meaningfulness has been argued as a positive or negative experience that is significant to an individual based on their psychological needs.

This brief discussion of some elements of experience sets a stage for a discussion of the main concepts involved in this study. It is perhaps a long journey from Dewey to Hassenzahl, and some of the intermediate steps will be taken in the next sections, before returning to this discussion towards the end of the chapter.

Attributes of Seniors Affecting their Access to Online Services

More than 14% of Canadians report having a disability, with this figure increasing to 43.2% for Canadians over the age of 65. The trend in these numbers is upwards, due to the aging of the Canadian population and the increased percentage of seniors relative to the total population (HRSDC, 2009). Coincident with the increase in aging and disability reporting is the proliferation of service access being shifted to information and communication technologies and internet websites (Sayago & Blat, 2010). However, designs need to be mindful of how cognitive, physical, and memory disabilities can affect the ability to access websites (Kurniawan & Zaphiris, 2005). The large percentage of seniors wih disabilities, along with the aging of the

Canadian population, necessitates accessible websites. Seniors need access to services that are increasingly offered online through government websites.

The attributes of seniors are discussed widely in the literature, and while a decline in motor skills is acknowledged to be the case as we age, it is the matters of cognition and perception that are most often discussed. Kurniawan and Zaphiris (2005) claim that the most discussed disability for web accessibility is that of visual impairment (perception), but note that there can be a cumulative effect when motor and cognitive impairments also exist. For example, a decline in motor skills can cause older users to have difficulties using keyboards and mouse (Hanson, 2011). In this section I provide a short overview of the research surrounding cognition and perception issues, as these are the most discussed issues in the literature. These issues point towards the diversity of abilities in and potential limitations faced by the senior population (Dickinson, Eisma, & Gregor, 2011).

Despite some opposing views, it is widely suggested in the literature that human cognitive processes decline with age and that this decline affects the use of technologies by seniors (for example see Sayago, Sloan & Blat, 2011). In the literature on technology and aging, human cognition or intelligence is often discussed in terms of fluid and crystallized (e.g. see Czaja et al., 2006, Hanson, 2011, Sayago et al., 2011). Fluid intelligence represents logical thinking and problem solving abilities, while crystallized intelligence represents the abilities to use knowledge and experience (Burmeister, 2010). There is some agreement that fluid intelligence undergoes a decline with age, while crystallized intelligence does not (Hanson, 2011). Fluid and crystallized intelligences have been found to be important predictors of technology usage (Czaja et al., 2006), and there is considerable discussion and theorizing that an abundance of crystallized intelligence ability can compensate for a reduction in fluid

intelligence, as one ages (Czaja et al., 2010). For example, Sayago et al. (2011) concluded that older persons use and develop strategies (e.g. note-taking) to cope with declining cognitive abilities.

It is important, however, to be wary of sweeping generalizations about seniors. Dickinson et al. (2011) found that it is difficult to distinguish conceptual barriers to web access from cognitive difficulties experienced in accessing the web. Cognitive difficulties might include age-related memory and processing difficulties, while conceptual barriers might include difficulties with language and jargon (Dickinson et al., 2007). The finding of Dickinson et al. (2011) matches that of a previous study conducted by Fairweather (2008) who found no significant effect of age or experience on study participants' success rates with an assigned web-based problem. In this study it was found that seniors followed different paths to problem solving than did younger adults, but with similar success rates. This finding led Fairweather (2008) to conclude that different types of users may behave differently, and to recommend that assumptions about user groups should be avoided. The finding also rejected the notion that older users address problems in the same way as younger users, but with less efficiency, speed, and accuracy. Instead, Fairweather (2008) suggests that older users may benefit by their previous experience, such that their behaviours become somewhat automated or routine through practice, thereby reducing the work required by their cognition and, in this sense, compensating for any possible age-related cognitive declines. These disagreements in the literature demonstrate at least two broad-based positions on the existence or non-existence of age-related cognitive decline, and my position in this study is that I will not address these issues directly through cognitive measures, but am mindful of the arguments.

As noted earlier, issues arise about the abilities of older people to adapt and use new technology (Hanson, 2011). Wandke, Sengpiel, and Soriksen (2012) in discussing this subject, note six myths or stereotypes about older people using computers. The myths that they have defined in their argument are listed as (p.564):

1. Future generations of older people will use computers without problems…
2. Older people are not interested in using computers…
3. Older people consider computers as useless and unnecessary…
4. Older people lack the physical capabilities to use ICT
5. Older people simply cannot understand interactive computer technology
6. You can't teach an old dog new tricks.

In their discussion, they note a grain of truth in each myth, while arguing that "…the myths are improperly overgeneralized and, therefore, often wrong. Such myths are problematic" (p.564). The argument concludes that, if left unchallenged, these myths become reality, to the detriment of older people. In arguing against the first myth, like Hanson (2011), the authors note the progressive nature of the development of technology. The suggestion here is that future cohorts of older people will struggle with technology, as the technology changes and new difficulties with the new technologies will emerge. As the authors state, "…this myth is based on the assumption of a singular cohort effect which will dissolve over time. In fact, this effect is continuously renewed…The task of getting people to use new technology remains" (p. 655).

In a like manner, Wandke et al. (2012) argue against the other five myths. They point out that '…learning is different in old age…" (p.569), and that, "…older people should receive special attention in regard to design and support" (p.569). They conclude by noting that

information and communication technologies "…offer vast potential to improve the quality of life of older people" (p.570). They describe the barriers to older peoples' usage of this technology as:

> These barriers can be described as myths which affect folk psychology and public opinion. Although these myths contain a grain of truth, we argue that they are unfortunately overgeneralized, as there are many empirical studies which have revealed these myths to be completely or partly wrong. They must be confronted with facts, as myths have a tendency to be self-fulfilling and self-reinforcing.

In this line of argument, the barriers to technology usage by older persons can be thought of as self-imposed or societal-imposed due to misconceptions (myths) about their abilities. As Wandke et al. (2012) challenge us, it is imperative to look to the facts rather than be caught in the web of myths.

In Silva, Braga, and Teixeira (2014) we find an overview of recent studies on human-computer interaction and age. Their review notes a number of reported age-related characteristics, including decline in vision, strength, memory etc. One conclusion of their review, based on the extant literature, is that designing for one age group and then adapting the design to another age group is not the optimal approach. In the context of older persons, this approach serves to attempt to modify an existing technology "…to cope with age related limitations instead of considering the broader context and a possible different approach to the problem from the start" (p.186). This has the undesired effect of relying on preconceived ideas about the characteristics of older persons, and the authors make note, similar to Wandke et al. (2012), of the erroneous nature of these preconceived ideas. For example, they assert that "Ageing does not

always translate to cognitive disadvantage" (p.186), noting that certain advantages of ageing could be emphasized in the design process rather than focusing on perceived disadvantages. In other words, design specifically for the older user, and don't be caught in preconceived notions about their declining abilities.

An earlier review of this topic maintains that with older users using computers more often, it is imperative to continue to study their usage as their needs are different from other age groups. Wagner, Hassanein, and Head (2010) examined 151 articles on the topic of computer use by older adults. Among other things, the authors noted that: "The most common use of computers and the internet for older adults appears to be for communication and social support" (p.873). A barrier identified to the use of computers by older adults is the lack of perceived benefit to using them, an attitudinal issue on behalf of the older adults, and something the authors feel can be addressed in order to encourage usage. This study takes a similar path to Silva et al. (2014) in recommending the consideration of older persons' physical limitations, while encouraging a multi-disciplinary approach and emphasizing the barriers that older adults have placed upon themselves – particularly the perceived lack of benefits and lack of motivation to use computers.

A study of eye-tracking and usability performance found that there was a difference in how younger and older users looked at monitor screens when performing website tasks, and that older users took longer to perform tasks than younger users. Romano Bergstrom, Olmsted-Hawala, and Jans (2013) gathered the data from five previous independent studies using eye-tracking to monitor website navigation. They found that there were differences in the areas of interest that older and younger users viewed, and that younger users demonstrated higher efficiency and accuracy in their navigation. They observed that older adults tended to view the

centre of the screen and not peripheral information, putting them at a disadvantage and resulting in longer task times than younger users. They conclude (p.547) that:

> Website design should be centered on the end user, including older adult users, so that they can access information and maintain independence with as much ease and satisfaction as younger adults do. It is important to consider issues associated with aging as design teams develop and modify websites.

These recent studies all suggest that attention should be given to the needs and characteristics specific to the older user, without falling prey to inaccurate myths about said needs.

Accessibility and Usability of Online Services by Seniors

A fundamental principle of the World Wide Web Consortium (W3C) is to ensure the inclusion of all humans in the accessing of, and the benefits incurred by, using the internet. The Web Accessibility Initiative (WAI) of the W3C produces guidelines and resources to make the Web accessible for people with disabilities. The guidelines include web content accessibility guidelines (WCAG), and the most recent version (2.0) was released as a recommendation in 2008. In addition to making the Web more generally accessible to people with disabilities, the guidelines are also intended to assist in making the Web more usable by older persons with age-related ability changes and to improve web usability in general (W3C WCAG 2.0, 2008).

The introduction section in the WCAG 2.0 guidelines notes that there are at least two major goals involved in the accessibility initiative. The first goal is ensuring accessibility to the Web by persons with disabilities. The second goal is improving the usability of the Web for older users and users in general. Since people often develop impaired perceptual, physical, and cognitive abilities as they age, these goals overlap. A premise of the guidelines is that websites

designed to be more accessible to persons with disabilities will also be more accessible to older persons (W3C AGE, 2010). This particular premise has been challenged by researchers who feel that guidelines should target user groups more specifically (for examples see Money, Lines, Fernando & Elliman, 2011). Some researchers feel that WCAG should include success criteria that specifically targets older users (Affonso de Lara et al., 2010).

In the context of WCAG 2.0, accessibility is governed by a set of principles that determine whether persons with disabilities can access the web. These four principles are that web content must be perceivable, operable, understandable, and robust. Guidelines are derived from the principles, and success criteria detailed in the guidelines provide measurable characteristics of web content (W3C WCAG 2.0, 2008). This process suggests that accessibility can be reduced to a set of characteristics of a web artefact. Romen and Svanaes (2012) debate this latter point, arguing that applying WCAG is not sufficient to guarantee the accessibility of a website. Their research showed that using WCAG as a heuristic to determine website accessibility only accounted for half of the accessibility problems encountered by the persons with disabilities in their study. It is their suggestion that the W3C WAI expand the definition of accessibility to include the ISO 9241 criterion that usability is achieved when persons with disabilities can achieve specific goals upon usage of a website. Their argument is centred on this expanded view of goal achievement in the use of a website. In their view, goal achievement focuses accessibility measurement onto the user experience instead of merely measuring websites against WCAG standards.

Accessibility characteristics might be considered as a subset of usability characteristics of a digital artefact. Usability cannot be measured, except by its absence – one can only measure how unusable something is, by measuring the problems people experience using it (Rubin and

Chisnell (2008). Accessibility can be measured by analyzing a website for conformance to the success criterion of WCAG (W3C WAI, 2005b). By measuring accessibility (in WCAG terms), we are measuring a subset of usability. We are not measuring the full user experience, particularly the experience of accomplishing a goal (Rubin & Chisnell, 2008).

What is the purpose of accessibility? Is it merely to have access to web content? Or is it being able to use web content in the achievement of specific goals? WCAG's accessibility definition does not define accessibility explicitly (Romen & Svanaes, 2012), but it does define its purpose as that of enabling persons with disabilities to access the web. WCAG's definition may be surmised by the principles it uses – accessibility means that web content is perceivable, operable, understandable, and robust. These principles don't explicitly demand that a user be able to achieve goals, but one could argue that implicitly they suggest that a user can use the web content to achieve their goals, in concert with Romen and Svanaes' (2012) argument.

Consider the hypothesis that accessibility is necessary but not sufficient for a website to be usable. This hypothesis initially begs the question: accessible for whom? Clearly, the objective of WCAG as defined in its purpose statement is to make the web accessible for persons with disabilities. However, usability might be another matter. In Table 1 the concepts of accessibility (WCAG) and usability are considered against the perceived requirements of persons with disabilities (PWD) and persons without disabilities (PWOD):

Table 1.

Accessibility and Usability Requirements for PWD and PWOD

Group	**Accessibility (WCAG)**	**Usability**
PWD	Deemed to be necessary for PWD to use technology	Receives greater consideration once accessibility is achieved
PWOD	Possibly not as necessary for PWOD to use technology	Possibly an initial focus of design for this group

In this table I am suggesting the possibility of a difference between the user needs of PWD and PWOD, while acknowledging that the argument that following accessibility guidelines improve website usability for all (W3C WAI, 2008). PWD *require* accessibility standards whereas this may not be as likely true for PWOD. Some recent research suggests that standardized conformance to WCAG accessibility guidelines is insufficient to fully address the user needs of individuals (Cooper, Sloan, Kelly & Lewthwaite, 2012; Romen & Svanaes, 2012). Cooper et al. (2012) suggest that user needs be prioritized over artefact accessibility testing. It should be noted, however, that WAI guidelines including WCAG 2.0 are based on input from various stakeholders including users (W3C WAI, 2008). In this approach, user needs are considered during the design of the guidelines.

Recommended guidelines such as WCAG 2.0 are amenable to artefact accessibility testing of the design parameters. The argument could be made that more attention should be focussed on user needs during the development of accessibility guidelines, if artefact accessibility testing is deemed insufficient in promoting the accessibility of websites. Artefact accessibility testing is a part of my study and will be discussed in greater detail in the

methodology chapter. The results of artefact testing compared with observations and user reports of experiences with websites helps further the debate on this matter of the prioritization of user needs over artefact testing.

Measuring accessibility.

Rubin & Chisnell (2008) consider accessibility and usability to be siblings, in the sense that accessibility is what makes a website usable to persons with disabilities. The World-Wide Consortium (W3C) defines web accessibility as meaning: "…that people with disabilities can use the Web" (W3C WAI, 2005a). These two approaches to accessibility essentially frame the concept of web-accessibility as an enabler for persons with disabilities (PWD), enabling them to make use of the web. To enable this usage, web accessibility can be thought of as being operationalized as the measurement of the removal of web-design artefacts that presents barriers for PWD. This removal of accessibility barriers is accomplished through the conformance to a set of web content accessibility guidelines (WCAG) produced through consultation with, among other groups, PWD that use web resources. Consequently, the W3C web accessibility initiative, through its focused approach to and operationalization of the construct of accessibility, has assisted in the ability to measure that construct.

However, for non-disabled persons, accessibility (as defined in this approach) may not be as necessary to their usability of the Web. Accessibility is a special case of usability linked to disabled persons and is a necessary pre-requisite to their web usage. For a disabled person, then, accessibility must come first, before further usability can be considered. Accessibility is a necessary condition for usability by disabled persons, but is it sufficient to make websites usable?

One school of thought considers that websites should not only be universally accessible but also should satisfy user expectations (e.g. Subasi, Leitner, Hoeller, Geven & Tscheligi, 2011). In this conceptualization one might consider that accessibility is the primary piece that opens the door to meeting user expectations. We may think of usability as akin to meeting user expectations in terms of the five attributes of usefulness, efficiency, effectiveness, learnability, and satisfaction (Rubin & Chisnell, 2008). Usability could be considered to be a primary goal of website function and design, once accessibility has been provided.

From accessibility to usability.

Current accessibility literature suggests that not enough attention has been paid to the sibling concept of usability. Usability can be measured by user experience, by measuring the problems experienced by people in using a website or web-page (Rubin & Chisnell, 2008). So closely does usability seem related to accessibility that the W3C definition and operationalization may appear to *include* usability. However, the indicators (White, 2009) of the two constructs may in fact differ considerably.

For example, Cooper et al. (2012) argue in support of elevating the user experience above mere automated software accessibility testing metrics that measure properties of a website or web-page. It is their contention that web accessibility metrics (WAM) should not be considered as superseding the analysis and support of a positive and subjective user experience. Web accessibility is not intrinsic to a digital resource, and metrics tend to treat it as such, to the detriment of the consideration of the user experience (Cooper et al. 2012). Cooper et al. (2012) would rather see a relational approach to accessibility where the relation between the user and

the digital artifact is given precedence over WAM. It is their view that universal guidelines may produce counter-productive outcomes with a paternalistic tone.

The argument of Cooper et al. (2012) suggests that accessibility has become linked too closely with automated web accessibility testing metrics which minimizes the subjectivity and context of the user experience. This school of thought considers that websites should not only be universally accessible but also should satisfy the user expectations (e.g. Subasi et al. 2011). The importance of the user relationship to websites is also noted by Byun and Finnie (2011), who argue that government websites should satisfy users. User satisfaction is determined by the users' perception of the usability of a website, and can be measured by the users' desire to revisit the website (Byun & Finnie, 2011).

The idea presented here is that a user will revisit websites that they perceive to be usable. In Byun and Finnie's (2011) model, perceived usability determines user satisfaction. In this model, the concept of perception is expanded beyond the mere mechanical ability of the senses to access a digital artefact. The model demonstrates an expansive view of perception as entering into the area of cognition, suggesting, at the very least, that the two concepts of perception and cognition overlap, and cannot thus be fully separated from each other. The argument of Cooper et al. (2012) suggests a similar conclusion – adhering to WAM testing without further considering user experience elevates the mechanics of sensory perception above the experience of user cognition. Their suggestion is to change the focus from digital artefacts to the user, arguing for the relational approach to accessibility.

The relational approach suggested by Cooper et al. (2012) encompasses the analysis of user behaviour through the use of analytics that gather data about users' interactions with web

resources. These data can provide a means by which to understand problems that users experience with web resources, thereby providing a better understanding of the user experience. Understanding the user experience can help to determine what users want in addition to how they behave. In a similar approach, Newell (2008) argues that consideration should be given to what older users want, and not just what designers perceive them to need.

Jaspers (2008) provides us with an analysis of three methods for usability testing: heuristic evaluation, cognitive walkthrough, and think aloud. WCAG evaluation most resembles a heuristic evaluation, while Jaspers (2008) maintains that a combination of methods would be most effective in assessing usability. Her perspective is similar to Cooper et al. (2012) in that both recommend more expansive testing to assess usability, rather than relying solely on heuristic methods.

If accessibility is about making something available for use by PWD, what is usability? Usability is more than just making something easy to use. It also incorporates "…the ISO 9241-11 usability standard definition, which says that something is usable if it is effective, efficient and satisfying to use…" (Stewart, 2012, p.645). Numerous studies have been conducted to measure these three characteristics of usability. For example, Kincl and Stach (2012) examine the effect of perceived task performance on user satisfaction with selected websites. The goal of that study was to compare user ratings on six website satisfaction criteria (before and after performing a specific task on the website). The results showed an asymmetrical effect. Negative task performance resulted in a less positive assessment of the website afterwards, while positive task performance did not alter the perception the user had of the website before conducting the task. What this study demonstrated was that the specific user satisfaction criteria (colours,

navigation, layout, visual appeal, content, and overall impression) were judged more harshly if the user had difficulty in task performance. A significance of this finding might be that the ability to use the website is more important that the style of it. The researchers found that the most important of the criteria were navigation and content, suggesting a usefulness and usability emphasis as opposed to an aesthetic emphasis in the user judgements.

In Sørum, Andersen, and Vatrapu (2012) we find a somewhat different suggestion. This study measured website quality and user satisfaction. The study used websites that had previously been recognized for high-quality, based on quality criteria measured by heuristic evaluation and objective measures such as ability to change text size, download time, content, navigation structure, and the use of sitemaps. User satisfaction was measured using surveys about these websites. User satisfaction with a website was measured as ease of finding information, content, and usefulness. The user survey scores were compared to the website heuristic evaluation scores. The study found a low negative statistical correlation between the two scores, and suggested that the website heuristic evaluations were technical in nature, and tended to ignore the users in this evaluation. The authors concluded that there is a strong accessibility emphasis in the criteria in the heuristic evaluation, and that this emphasis does not correlate entirely with user satisfaction.

Early attempts to measure user website experiences in a quantitative fashion involved the use of such concepts as *lostness* (Smith, 1996). In this measure, a user is considered lost when they can't locate desired information that exists in the system. Lostness is quantified by assigning values to statements made by the user during testing, video examination of behaviour and counts including that of the numbers of websites accessed while searching. *Perceived disorientation* is a more recently used concept related to lostness, and has been measured as a

subjective variable through the use of survey responses (e.g. Chen, Lin, Yen, & Linn, 2011). The Chen et al. (2011) study examined the effects of content familiarity, media richness, and site breadth on user perceptions of disorientation, engagement, and intent to use the website again. It was found that the independent variables impacted the user perceptions in a synergistic fashion. The researchers conclude that weakness in one variable can be compensated for by strength in another. For example, increasing the media richness and including familiar content can reduce the tendency that a deep site with fewer hyperlinks per page has in generally increasing lostness. The significance of their findings is the conclusion that several factors interact to affect user perceived disorientation, and they suggest that these factors be studied together.

Hassenzahl and Tractinsky (2006) make note of what they term "…the potential mismatch between designers' intentions and users' actual experiences" (p.92). It is their contention that research on human-computer interaction (HCI) should emphasize the user experience (UX) instead of focusing primarily on task completion, as traditional usability studies have done. They think of UX as a multidimensional model that includes attempts to understand the role of affect on technology use. In its focus on the user, the authors claim that UX aims to assist in understanding how quality experiences arise as opposed to just focusing on removing barriers and problems. In their model, the focus of UX is a change to traditional views towards HCI. As Hassenzahl and Tractinsky (2006) state:

> UX in the sense of a *positive* HCI would, thus, focus on how to create outstanding quality experiences rather than merely preventing usability problems. Again, this will question another implicit assumption of traditional HCI, one that equates high quality with the absence of problems (p. 95).

This argument is supported by the findings of Sørum et al. (2012) noted above. In Sørum et al. (2012), high quality websites (as determined by heuristics designed to alleviate problems) did not positively correlate with user satisfaction, a result that challenges the implicit assumption of traditional HCI noted by Hassenzahl and Tractinsky (2006). If such high quality websites are not solely to be judged by the absence of problems, how then can they be judged? As Hassenzahl and Tractinsky (2006) note, evaluation can and should include the use of UX evaluation towards the goal of creating more than just an absence of problems. As they see it:

> But just as there is much more to wellbeing than the absence of malady, so must there by more to UX than the absence of problems. From our perspective, one of HCI's main objectives in the future is to contribute to our quality of life by designing for pleasure rather than for the absence of pain. UX is all about this idea. (p. 95).

In the next sub-section I will review how user experience can be measured. As we shall see, it contains elements of usability measurements and accessibility measurements, but the goal of UX research is to create positive user experiences (Hassenzahl & Tractinsky, 2006).

From usability to user experience.

Neilsen (2012a) attributes five components to usability: learnability, efficiency, memorability, errors, and satisfaction. Each component can be measured in terms of how it contributes to the ease of use of a website. The overriding concern is with how easy and pleasant a website's features are to use. His approach to usability has similarities to that of Hassenzahl and Tractinsky (2006) in emphasizing the pleasure component. Rubin and Chisnell (2008) claim that usability is composed of six elements: usefulness, efficiency, effectiveness, learnability, satisfaction, and accessibility. In their scheme, Neilsen's (2012a) memorability is merged with

learnability, and they have included usefulness and accessibility as components. However, usability studies often focus on performance metrics such as error rates, speed, learning, and goal achievement rates (Fisk, Rogers, Charness, Czaja, & Sharit, 2009), which represent only four of the six components (error rates, efficiency, learnability, and usefulness) in Rubin and Chisnell's (2008) model. The remaining components are satisfaction and accessibility.

User experience (UX) relies heavily on measures related to user satisfaction, in particular to the positive aspects of interaction outside the traditional task completion utility measures (Bargas-Avila & Hornbaek, 2011). User satisfaction is considered to be a subjective rating of a website by a user based on their feelings, perceptions, and opinions. These feelings, perceptions, and opinions can be gathered through questionnaires and interviews (Rubin & Chisnell, 2008). Interestingly, in publications about UX research, half use qualitative, 33% use quantitative, and 17% use mixed methods approaches to collecting data according to a review study by Bargas-Avila & Hornbaek, 2011. Their review study of the publications on UX research finds a large shift towards the qualitative methodologies accompanying the shift to UX research from traditional HCI usability studies.

Bargas-Avila & Hornbaek (2011) conclude that it is difficult to categorically associate UX with non task-oriented goals and usability with task-oriented goals. They claim that these goals are often inseparable and goals are not suited for differentiating between the two. This conclusion is different from that held by Hassenzahl and Tractinsky (2006) and has moved the conceptualization of UX towards viewing experience as containing new dimensions (e.g. engagement, enchantment) and being multi-dimensional in character. These researchers believe that UX research builds upon traditional usability research and concepts, while cautioning about the challenges faced in developing proper measures and definitions for newly conceived UX

dimensions of experience. More recently efforts are underway to map out the theoretical underpinnings of the field of user experience (for example see Obrist, Roto, Vermeeren, Vaananen-Vainio-Mattila, Law, and Kuutti, 2012).

If the unique measures of UX are in a developing stage, then is it possible, bearing in mind the suggestion that UX builds upon usability (Bargas-Avila & Hornbaek, 2011), to utilize usability measures while attempting to extend them into the dimensions of UX? I think that this is a worthy goal to pursue, helping us understand how UX might expand upon usability and enabling us to use this understanding to develop better websites for seniors. In this study I employ usability measures in the study of seniors accessing websites, while measuring qualities sometimes considered to be dimensions of UX.

Completing the circle - accessibility and user experience (UX).

Clarity of definition of key concepts used in research questions is of central importance in any study (White, 2009). In this study, I measured the UX of seniors while accessing government websites, and compared the UX to other variables identified in my research questions. A definition of UX is required, acknowledging the argument in the literature cited above that claims that such a definition is unclear and still under development. What is needed is a working definition of UX.

A fundamental question remains as to the ways UX extends from usability. Hassenzahl and Tractinsky (2006) have suggested that a differentiating aspect is a UX emphasis on affect as perceived by the user. Their suggestion is that "…non-instrumental needs must be better understood, defined and operationalised" (p.93). They consider instrumental needs to be those that have been tested in usability studies, for example - efficiency, task completion, etc. Non-

instrumental needs include the idea of hedonic quality, which they argue is relevant to fulfilling "…an underlying human need – a need for being stimulated, to perfect one's skills and knowledge, to grow" (p.93). This difference in emphasis served in the working definition of UX for the purposes of my study. I used traditional measures of usability in examining the user experiences of participants, while employing methods designed to discover the affects of the experiences that the users created. In other words, I included the affective experiences of the participants as perceived and reported by them while using the technology.

In order to arrive at a working understanding of the concept of affects for use in this study, I turn back to the concept of accessibility as previously discussed. Accessibility can be considered to be a technical term in the measurement of it, in the sense of objective measures of the lack of problems that a user encounters in using a website. The measure is obtained through evaluations of the website's characteristics against a pre-determined standard. Conversely, affects can be considered to be subjective measures, individually held and measured through evaluation of responses made by users. Affects then can include user satisfaction, as expressed through questionnaires and other responses, and other elements of affect including feelings, moods, and attitudes (Scherer, 2005).

As noted above, we were left with user satisfaction and accessibility as being generally under-utilized in usability studies (Fisk et al., 2009). An emphasis in my study will be on these two components of usability, with a view towards expanding upon the component of user satisfaction by attempting to capture data concerning the affects of the technology upon the study participants. Previous work in the measuring of usability recommends a research agenda that would include expanding upon the traditional measures and constructs of usability (e.g. Hassenzahl & Tractinsky, 2006; Hornbaek & Law, 2007; Hertzum & Clemmensen, 2012). There

are views in the literature that suggest that UX subsumes usability and/or is an enhancement of the usability measure of user satisfaction (Law, 2010).

Consequently, I measured the accessibility of websites and web-pages through the indicators of web-accessibility metrics as determined by their conformance to W3C guidelines, in order to give a sense of the environment in which the tasks are performed, and to check on the compliance of the websites with the guidelines. I measured the usability of websites and web-pages as indicated by the user experience of study participants through user observations, user narrative, and the performance of users on specified web search tasks.

Further discussion of the Conceptual Framework

This section continues the discussion of the theoretical and conceptual framework for my study. The concepts of learning and adapting help provide the context for a theory-based argument on user experience as a formative conceptual framework for my study.

Learning and adapting.

As a preliminary definition, learning in the context of this study can be thought of as including processes by which a web-user adapts to the web environment. It can be measured by the web-user's ability to achieve desired goals or assigned tasks using the web (see learning definition Nielsen, 2012a; also learnability Rubin & Chisnell, 2008),

Adaptability was seen in the Sayago et al. (2011) study. It was their finding that cognitive difficulties have a more severe impact on accessibility than visual limitations or mouse-use limitations. To overcome cognitive (i.e. memory) issues, seniors made notes to help them to remember the steps and areas they used in navigating the web, and used these notes to assist

them when re-navigating. Sayago et al. (2011) concluded that these adaptations were more time-consuming than the adaptations made to overcome visual and motor issues (e.g. mouse skills), claiming that the latter are overcome with experience coupled with a sense of inclusion in using technology. The ability to adapt arose from a feeling of being a competent web-user and a desire to be included in web-use.

Studies such as Sayago et al. (2011) suggest that the use of technology is related to the usefulness that seniors attribute to it, along with the confidence that seniors have in their ability to adapt. These factors are in turn related to the design of the technology (e.g. website) being used (Subasi et al, 2011). When seniors feel confident with technology, and less-alienated from technology, they develop their own inclusive strategies to overcome difficulties with internet technologies and websites (Sayago et al., 2011). In this manner, adaptations are related to barriers, perceived and otherwise.

Demonstrations of collaborative learning at work amongst older users have been observed. In Sayago and Blat's (2010) ethnographic study of older users and the web, it was reported that the social aspect was pronounced. The participants in this study acted together as groups to learn computing tasks. The groups functioned to support individuals in their learning, as well as to assist in memory and computer jargon issues. Participants reported that remembering how to use the computer was more important than text size; participants could adapt to text size by sitting closer or removing/wearing glasses, but forgetting how to navigate and use the computer presented larger problems. The use of other group members' memory helped to alleviate this problem.

Learning and adapting to internet websites while accessing online services involves a number of processes, as outlined above. While some seniors may experience a reduction in aspects of cognitive ability (e.g. memory) as they age, the loss may be compensated for by their experience and held knowledge (Czaja et al., 2010). Cognitive load theory provides further insight into these concepts.

Cognitive load theory.

Cognitive load theory (CLT) is relevant to the notion of age-related issues regarding memory. Age-related issues are a consideration in the aims of the W3C. In CLT, it is theorized that the human cognitive architecture is generally structured such that a limited capacity of working memory (WM) is available to a person (Sweller, Merrienboer, & Paas, 1998). This limitation means that a person can only hold a finite amount of information in their consciousness at any one time. As the conduit to long-term memory formation, the limited WM restricts the amount of information that can be processed. In CLT it is theorized that age-related memory decline may further restrict this existing limitation by reducing the WM capacity and thereby also reducing the processing capacity (Kirschner, P.A., 2002).

The finite amount of information capable of being held in consciousness at any one time is affected by three cognitive load factors (Sweller et al., 1998). In the context of learning materials, intrinsic load is a demand placed on WM by the nature of the materials themselves. In the same context, extraneous load is a result of a material's design that acts to increase memory processing (e.g. poorly designed material), and germane load is a result of a material's design that acts to facilitate memory processing (well-designed material). Some studies (e.g. Sayago & Blat, 2010) suggest that reducing the cognitive load will help seniors by offsetting decreased

memory capabilities. It is also possible that distributing the cognitive load over a group may result in a lowering of any one person's cognitive load to the advantage of the group's processing capacity (Zhang, Ayres, & Chan, 2011).

As argued previously, usability includes the concepts of error rates, efficiency, learnability, usefulness, satisfaction, and accessibility (Rubin and Chisnell, 2008). Two conceptual models are presented by Hollender et al. (2010) towards the goal of integrating CLT concepts with usability concepts within the field of research on educational software. The first of these models proposes that extraneous load can be reduced by proper attention to the design of software, with a particular emphasis on incorporating usability principles into CLT design research in order to increase germane load. The second model proposes the integration of CLT concepts into usability concepts in the design process.

Both of these models propose an integration of one set of concepts with the other, with attention given to which set of concepts is the departure point of the design process. In my study, I am arguing UX as the focus for evaluation of website usability, and have used CLT to help inform my understanding of the possible characteristics of the population I will study. I will not be measuring CLT concepts in my study. It is mentioned here to show a competing theoretical approach to increasing the usability of websites by older people.

Usability measures

As noted earlier, Rubin and Chisnell (2008) list the six elements of usability as usefulness, efficiency, effectiveness, learnability, satisfaction, and accessibility. These elements can also be considered to be reduced to a fundamental usability construct that includes the qualities of efficiency, effectiveness, and satisfaction (Sauro & Lewis, 2012). A number of

measures have been devised in order to address these latter elements, and thereby arrive at an evaluation of the usability of an item. Among the measures commonly collected are completion rates, task time, satisfaction, error rates, and problems (Sauro & Lewis, 2012). To these fundamental usability constructs, Brajnik and Giachin (2014) have reported on a number of other constructs which they consider to be user experience (UX) characteristics, noting the complexity of the overarching concept.

Brajnik and Giachin (2014) state that UX elements extending from usability are quite varied and contextual. As a result it is difficult to suggest standardization of metric usage when considering them for a study. It is their recommendation that measures be chosen and adjusted to suit the context, and that several different measurement constructs be considered. In their study they noted that, "Therefore, one could view UX as an umbrella concept that goes beyond usability and accessibility, embracing a range of properties that deal with many psychological, physiological and social human phenomena" (p.553). They list several other properties that they claim UX covers in addition to usability and accessibility. These aspects include aesthetics, emotions, perceived usability, hedonic attributes, cognitive load, interactivity, social responses, persuasion, and acceptability. In their study, they measured a subset of these attributes, having considered them to be the most appropriate to their study of age-related differences in the use of a technological device (digital thermostat). Their study collected both quantitative and qualitative data in addressing these UX attributes (and the common usability elements). The importance of the Brajnik and Giachin (2014) study to my work is in the extensive listing of UX attributes, the description of these attributes, and the suggestion to choose which ones to measure based on the context of research, in addition to the use of common usability measures. It is also important in highlighting the need to collect more than just objective data. As the authors state, "running a

usability investigation focusing on measurable objective data only is likely to miss important differences bearing upon acceptability and actual usage" (p. 562-3).

Other literature also recommends a combination of measurements in assessing usability, relying on more than just heuristics (e.g. Jaspers, 2000; Cooper et al., 2012). There is an argument to emphasize user experience instead of just traditional usability measures (Hassenzahl & Tractinsky, 2006). A number of studies have proposed using differing sets of guidelines in order to address older users' needs directly.

One earlier study used both a heuristic evaluation along with a usability study (Hart, Chaparro, and Halcomb, 2008). The heuristic evaluation had four expert evaluators assess a list of websites against a set of usability guidelines created by the United States National Institute on Aging and the National Library of Medicine. The websites chosen for assessment were ones that targeted older users. The guideline assessment produced compliance scores, and thirty-six websites were scored in this method. Three sites from the thirty-six were chosen for testing in a subsequent usability study, based on differences in compliance scores. The chosen sites were ranked as most, medium, and least compliant.

In the usability portion of this same study, twenty-one participants over the age of fifty performed five search tasks on each website. Their performance measures included completion rates, time, number of pages used, and user satisfaction measured by the SUS scale. The study found that the most compliant website had the higher task success, but the medium compliant website was more preferred by the users and had the highest System Usability Scale (SUS) scores (Brooke, 1996). The authors concluded that:

> The results of this study demonstrate the importance of using usability testing to evaluate a website in addition to its adherence to guidelines…Guidelines alone are not sufficient for insuring usability because they are often very general and lack the specificity necessary for detailed design decisions (p.198).

In a similar manner, I have chosen a combining of both guideline testing and usability testing as part of the method for my study.

Norval, Arnott, and Hanson (2014) maintain that a digital divide exists between younger and older users of social networking sites (SNS). Their study sought to investigate the barriers that these sites have for older users. In the first part of their study, they used focus groups of older SNS users to create a list of recommendations for these sites. In the subsequent part of the study, two different sites were evaluated by thirty participants. One site functioned as a control site, with the other site having been altered by employing the recommendations created by the focus groups. Measures included completion rates and SUS scores.

Norval et al. (2014) characterize their study as "a case study into the impact of the recommendations on a theoretical but representative SNS site" (p.3931). They found that the recommendations helped to improve the usability and task completion rate. It was among their conclusions that prior assumptions about the meanings of icons and other features should be reconsidered in terms of the older user, as they may present barriers to their use. In order to reduce the digital divide, the authors conclude that user studies with older users are important in producing design recommendations.

Another example of the use of heuristic testing is given in Patsoule and Koutsabasis (2012). This case study examines the redesign of a tourism website using a set of seven

principles and forty-five guidelines specific to web design for older persons, and a subsequent usability study involving twelve older users. The study stresses "…the importance of using both guidelines and usability testing when designing websites for older adults…" (p.563). Using guidelines amalgamated from previous studies, the authors employed evaluators to verify the validity of the set and then assessed a target website against them. The target website was redesigned to reflect the guidelines and then the two websites were compared in the usability study. The metrics used in the comparison included standard usability measures - task success, time, errors, and efficiency. The redesigned website was found to be more usable and preferable to the original website. In conclusion, the authors maintain that website design should be user-centred and include participation by older adults.

In Lynch, Schwerha, and Johanson (2013) we are presented with another study that developed a heuristic for the evaluation of website use by older adults. In this case the heuristic was developed by choosing guidelines from multiple sources, and surveying their efficacy with a group of older adults in order to ascribe a weighting to the individual guidelines. The resulting heuristic was used by two experts to evaluate twelve websites, providing a usability index for each website. Thirty-one older participants were employed in a subsequent usability study to evaluate three of the twelve websites. The three websites were identified as having low, medium, and high usability scores from the expert evaluations. Findings showed that heuristic evaluation was predictive in assessing websites and their subsequent performance attributes. It was also noted that the participants wanted a contact number on the website to reach out for further information and assistance. The authors concluded that their study has: "a) established a method of quantitatively evaluating a website's usability for older users and b) provided insight into conducting usability tests on various websites designed for older users" (p.416). They

recommend that more research be conducted using the System Usability Scale (SUS) and comparing the scores with task performance.

Chou, Lai, and Liu (2013) studied elderly people using Facebook. Their case study was conducted in Taiwan, and while they suggest that their findings may be limited to the specific living conditions and environment in that country, the idea that user studies are context-specific is in agreement with other studies noted above. Chou et al. (2013) call their method "an anthropological user centred approach, to explore the verbal behaviour of senior citizens while they accessed Facebook" (p.920). They warn of "the emergence of the disadvantaged elderly group of people, and…the problem of an information gap that exists among the elderly group today" (p.921). Their study findings draw attention to the idea that a user-centred approach to design may help to alleviate the information gap problems of the marginalized elderly.

User experience.

As noted above, Cooper et al. (2012) assert the necessity to prioritize user needs over focusing on the accessibility metrics of artefacts such as websites. As they state,

> This paper argues that web accessibility is not an intrinsic characteristic of a digital resource, but is determined by complex political, social and other contextual factors, as well as technical aspects which are the focus of WAI standardisation activities. It can therefore be inappropriate to develop legislation or focus on metrics only associated with properties of the resource. (Cooper et al., 2012, p. 1)

The authors further argue that an alternate route to providing inclusive online services is to support a user focus in online service design that works towards fostering a positive user experience during online service use. In their view, the concept of online inclusion should

expand from an objective artefact accessibility focus to a user experience focus that encompasses subjective and context-specific user perspectives on their usage of online services.

In the context of interactive technology, experience has been modeled by Hassenzahl (2008) as ongoing self-talk about human –product interaction. The self-talk includes the element of feeling, which constitutes a way of subjectively evaluating a product. As he states:

> Consequently, I define UX as a *momentary, primarily evaluative feeling (good-bad) while interacting with a product or service.* By that, UX shifts attention from the product and materials (i.e., content, function, presentation, interaction) to humans and feelings – the subjective side of product use. (Hassenzahl, 2008, p. 2)

By raising the profile of feeling within the human-product interaction, UX, in Hassenzahl's theory, focuses on a concept he calls hedonic quality. This quality refers to the perceived ability of a product to support the user in the achievement of feelings such as competence, autonomy, stimulation, etc. He calls these *be-goals*, as in the sense of being competent, being autonomous, and being stimulated. He contrasts these with *do-goals*, which he characterizes as goals found in traditional usability thinking that relate to the achievement of tasks. These task-related goals have a pragmatic quality, focusing on the usability of the product and supporting the fulfilment of be-goals. In this view of UX, the focus of human-product interaction is on feelings of well-being and not performance, thus emphasizing the usability concept of user satisfaction.

Returning to Dewey, it is helpful to note another similarity between his ideas and those of Hassenzahl. A connection between Dewey and Hassenzahl might be found in their use of the word 'quality' in their discussions on the criteria of experience. Where Dewey claims that experience is continuous, modifying the quality of previous experiences, Hassenzahl claims that

the examination of user experience involves the goal of contributing to the "…quality of life by designing for pleasure rather than the absence of pain" (Hassenzahl & Tractinsky, 2006, p.95). Hassenzahl (2014, p.8) posits that the pleasure/pain evaluation "…is important in shaping future behaviour…", while Dewey (1938/1997, p.27) states that "The quality of any experience has two aspects. There is an immediate aspect of agreeableness or disagreeableness, and there is its influence upon later experiences".

I suggest that Dewey and Hassenzahl are saying some similar things about experience. This chart summarizes what I think they are saying:

Table 2.

A comparison of Dewey and Hassenzahl's thoughts on experience.

	Interaction	**Continuity**	**Meaningful**
Dewey (1938/1997)	-situational, between individual and the (educational) environment they are in	-present experience incorporates the past and affects the future through created attitudes	-agreeableness – disagreeableness (quality of experience) influences future experiences
Hassenzahl (2014)	-situational between user and digital environment (device)	-memories of the past are modified in the present to create stories that affect the future	-pleasure/pain evaluation (quality of experience) shapes future behaviour

As noted previously, Hassenzahl et al. (2013) have argued that: "…it is actually the fulfillment (or frustration) of psychological needs that renders an experience positive (or negative) and personally significant, that is, meaningful". This conceptualization of meaningful has been used in the above table, and I argue that both Dewey and Hassenzahl et al. are making a similar point – that future behaviour/experiences are shaped by a quality of experience evaluation that is based on an individual's needs (psychological or otherwise). Where Dewey (1934/1997) speaks of the miseducative experience (as harmful to the growth of the individual), Hassenzahl (2014) speaks to the positive effects to be derived from attention to the creation of positive experiences.

As a consequence of finding similarities in their arguments, I have used their terms to augment the conceptual framework for this study, merging and overlapping them as appropriate. In the continuity box of Table 2 I have shown Dewey linked to the term attitudes and Hassenzahl linked to the term stories. I suggest that the conclusions in this box are similar, and that there are overlapping elements in the concepts of attitudes and stories. Attitude is defined as a way of thinking and feeling about something (Merriam-Webster (2014a). Is it useful to think of narratives as the expression of attitudes? Does narrative as a form of thought have an attitudinal component? Does narrative as a form of discourse express attitudes? These questions are prompted from my thoughts about the similarities between the two authors, and the similarities between the authors help inform the conceptual framework in this study.

One method to resolve the questions is to look at what a story (narrative) represents. Bruner (1987) tells us that one distinct form of thought and representation is that of narrative, prefacing his argument by maintaining that logical thought is not the only form. In describing the narrative form of thought, Bruner (1991) maintains that narrative "…operates as an instrument of

mind in the construction of reality" (p.6). Bruner (1996) notes that the behaviour of people in narratives is motivated by belief, values etc., suggesting to me a connection between narrative and attitude, when one thinks of the concept of attitude as being closely related to the concept of belief.

Recent user experience papers note that narrative discourse (in text form) can be representative of thought and feeling, or attitude. For example, Meneweger, Wurhofer, Obrist, Beck and Tscheligi (2014) have developed a framework to use in the classification of narrative gathered during user experience research. Four text types are described, with two being considered of most interest for user experience research - situative and evaluative text. Situative text represents process, context, and physical environment. Evaluative text represents thought, feeling, and impression. The use of narrative in understanding experience is a feature of narrative inquiry. As Clandinin and Connelly (2000, p.17) note, "For us, Life – as we come to it and as it comes to others – is filled with narrative fragments, enacted in storied moments of time and space, and reflected upon and understood in terms of narrative unities and discontinuities". It is their contention that "…narrative inquiry is a way of understanding experience" (p.20), and that narrative has an explanatory quality (Connelly & Clandinin, 1990). Additionally, these authors think of "narrative as both phenomena under study *and* method of study" (Clandinin & Connelly, 1990, p.4). As discussed further in the next chapter, this study gathered the two types of narrative texts of interest for studying user experience and analyzed them in terms of the experiences they represent.

Research Questions

I have designed my research questions to address the satisfaction and feelings of users, along with barriers and adaptations made, in relation to older users' experience in the environment of accessible government websites. My research questions centre on exploring the experience of the participants as they navigate these websites in the performance of information seeking tasks. The research questions are:

1. What is the user experience of seniors when using government websites tested for compliance with accessibility standards?
    a. What barriers, if any, do seniors experience when using government websites?
    b. How do seniors adapt to the environment presented by government websites while using them?
2. What are the relationships among the user experience of seniors, barriers experienced, adaptations made (when using government websites), and the reported accessibility rating (of the government websites)?

Question one broadly asks about the user experience of seniors within the contextual environment noted. As such, it relies on the concepts found in the user experience literature, as it builds on usability research. The two sub-questions address the specifics of barriers and adaptations, as experienced and undertaken by seniors in the noted environment. The second question seeks to explore the relationships between various conceptual elements, measured as noted in the following chapter.

Summary

In summarizing this chapter, I argued for the position that accessibility initiatives such as conformance guidelines should give a higher degree of consideration to the user experience of

the user groups being addressed. Several of the studies cited in this chapter have reported that issues other than those addressed by conformance guidelines are more important to seniors' needs (e.g., memory issues). In the study of accessible government website usage by seniors, it would be valuable to know whether the same issues exist – are seniors concerned with text size and other accessibility conformance-type criteria, or are there other issues they value more or are more applicable to their group? A way to address these questions is through examining the user experience of the seniors while using these websites. I have argued that a study of UX should focus on satisfaction issues, specifically including the feelings of users about the websites they are using. Consequently, I have created research questions that seek to address these issues. In the next chapter I review the research methods and design by which these questions will be addressed.

Chapter 3 - Research Design and Methods

Introduction

In this chapter I introduce and discuss the research design and methods that I used in my study. The methods for analysis of the collected data are reviewed with reference to the research questions. Relevant literature is reviewed in order to support a conceptual context within which the study is situated. The context is that of user experience (UX), defined as an application of usability concepts with the inclusion of the experienced affects such as feelings and emotions. This context is used to produce a framework for measuring and analyzing data collected during the study.

I used a mixed methods approach to collecting and analyzing data. The chapter includes discussion of both the quantitative and qualitative data collection and analysis strands and a section describing the pilot study wherein the data collection instruments were tested. It concludes with a note on the limitations, ethics review process, and the timetable of the study.

Conceptual Framework

Law (2011) states that: "Apparently, UX people tend to embrace eclecticism with no single paradigm or set of assumptions being rigidly followed, but drawing upon multiple theoretical approaches to gain complementary insights into UX" (p. 5). This suggests a contextual approach to theory usage in UX studies, a suggestion that Law promotes. In order to address the needs of seniors, a conceptual framework that used UX concepts was adopted along with standard usability measures, emphasizing a user-focused framework as discussed in the previous chapter.

Methodology of the Study

I used a mixed methods approach (Teddlie & Tashakkori, 2009) that enabled the collection and analysis of both quantitative and qualitative data. I selected a mixed methods approach for the study because it provides for multiple data sources and analyses, obtaining more evidence in the study of a problem (Creswell & Plano Clark, 2011).

Mixed methods research is considered to be a research paradigm, and it can be positioned as being between the extremes of purely quantitative and qualitative research (Johnson, Onwuegbuzie, & Turner, 2007). As Johnson et al. (2007, p. 113) describe it:

> Today, the primary philosophy of mixed research is that of pragmatism. Mixed methods research is, generally speaking, an approach to knowledge (theory and practice) that attempts to consider multiple viewpoints, perspectives, positions, and standpoints (always including the standpoints of qualitative and quantitative research).

This multiple perspective approach has met with challenges to its use, in what is known as the *incompatibility thesis* (Teddlie & Tashakkori, 2009). This thesis claims that integrating quantitative and qualitative methods is impossible as the underlying paradigms are incompatible. Counter to this position is the compatibility thesis, which rejects the either/or position and replaces it with a both/and proposition (Teddlie & Tashakkori, 2009). The philosophy of pragmatism serves to enable the mixing of quantitative and qualitative methods by rejecting a dualistic philosophy in favour of a pluralistic one. As Teddlie and Tashakkori (2009, p. 73) state:

> A major reason that pragmatism is the philosophical partner for MM is that it rejects the either-or choices from the constructivism-positivism debate. Pragmatism offers a third choice that embraces superordinate ideas gleaned through consideration of perspectives

> from both sides of the paradigms debate in interaction with the research question and real-world circumstances.

This philosophy is one that I am able to agree with. Pragmatism speaks to the value of utility, e.g. how well does theory work and/or how practical is the solution proposed? This utility of purpose matches well with my objectives in this study.

When choosing how to address the research questions I used, I decided that mixed data collection methods, and a mixed and integrated analysis of the data, provides the best approach. The constructs that I measured include adaptations, barriers experienced, accessibility, and user experience. Accessibility was measured quantitatively by automated software programs that produce a numerical count of compliance with WCAG 2.0 standards. Adaptations and barriers experienced were measured by observations and/or verbal utterances made by study participants. The making of researchers' notes while observing participants helped produce both quantitative and qualitative data.

Experience is often communicated through narrative (Bruner, 1987), and inquiry into experience helped reveal the perspectives of study participants. I analyzed the construct of user experience in two ways. A usability survey produced quantitative data. Semi-structured interviews along with participant observation provided qualitative data. It is my belief that gathering data in different forms (numeric and narrative) provided for a deep and broad understanding of my topic, while permitting the comparison and triangulation of the data during the integration phase. A deep and broad understanding and corroboration defines the mixed methods research approach (Johnson et al. 2007).

Research Design

The research design that I have chosen most closely resembles a parallel mixed design (Teddlie & Tashakkori, 2009). A parallel mixed design provides for both quantitative (QUAN) and qualitative (QUAL) strands in data collection and analysis. An example of this type of design is given in an article by Sayago, Guijarro, and Blat (2012). The article is a report on a study of internet web forms as used by older persons. Following a brief training period on how to fill out web forms, eighty-eight older persons (ages 65-75) were randomly divided into twelve groups. The groups were then instructed to fill out web forms requiring them to either create a Yahoo! or Hotmail e-mail account, or to book a flight with Vueling airlines. Each of these tasks presented a web form that used different displays to indicate required and optional fields in the form. The four different display types for the required fields were: large asterisks, standard asterisks, textual labels, and binary classification. The last type divided required and optional fields into two distinct sections, whereas the previous three types mixed the fields together, with the required fields being denoted by asterisks or text labels.

The Sayago et al. (2012) study reported that the ways of denoting fields on web forms had a statistically significant effect on the errors made on the forms. Errors were defined as the number of required fields not filled. Failure to complete a required field on a web form required the participants to make further attempts. The study was experimental in nature, in that the independent variable was manipulated through four field types, with the dependent variable being the error number. Participant interviews followed the experiment, and Sayago et al. (2012) used the analysis of the interviews along with participant observations to explain the results of the experiment. The experiment showed that the binary classification methods produced the

fewest errors, and the interviews maintained that binary classification was a "…familiar way for them to organise elements in the real world" (p.178).

Also of interest and use to me is the notion that this study represents "…a specific instance of the wider problem of reducing cognitive load" (Sayago et al., 2012, p. 173). The concept of cognitive load is relevant to my proposed research, and this study is useful to me in showing a manner in which the concept is considered relevant to the results of an empirical study. Additionally, the study measures the errors made during the performance of web form completion and does not measure the time taken. This approach is interesting to me because it differs from several other studies that measure task-time. The rationale given in this study is that time is not a useful parameter for older computer users as they are more interested in accuracy than time. The authors base this conclusion on their previous studies and the interview results from this study. However, the argument could be made that time is an important indicator of the degree of difficulty in achieving accuracy, and the authors do note that it would have been useful to know the relationship between speed and error number. Accordingly, I will build in a way to measure this relationship in my study.

My interpretation of the design of the Savago study is shown in Figure 1:

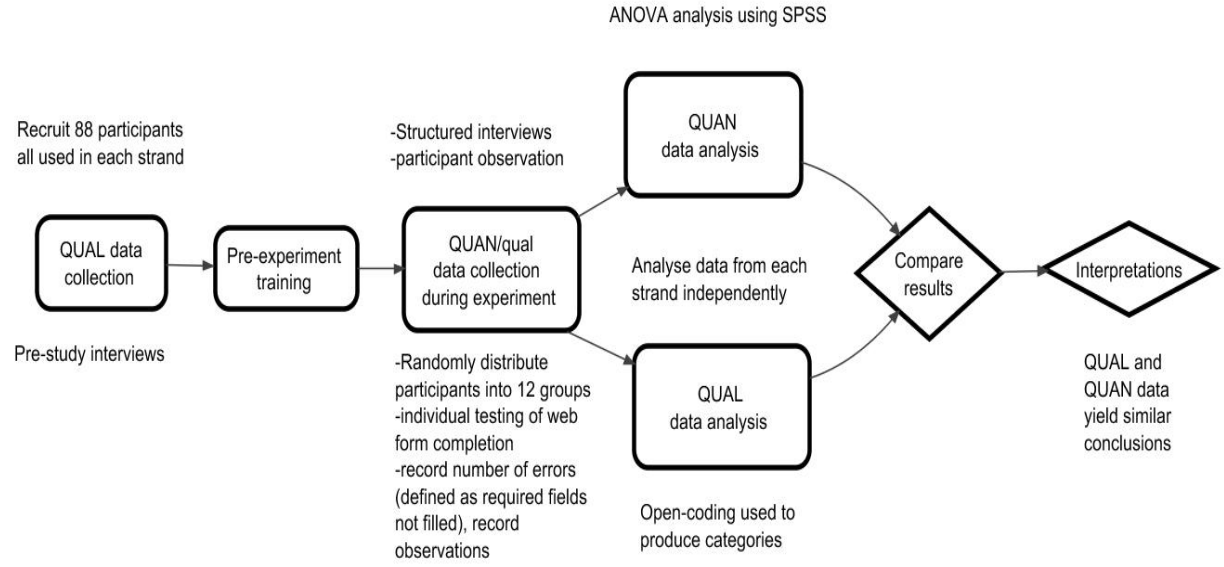

Figure 1. A variant of the parallel mixed design, derived from Sayago et al. (2012).

In the design pictured in Figure 1, quantitative and qualitative data collection occurs in a parallel manner, and analysis occurs independently during the same phase of the study as well (Teddlie & Tashakkori, 2009). The addition of an experimental component within the Sayago et al. (2012) study might also render its classification to be considered as a variant of the embedded experimental design model (Creswell & Plano Clark, 2011). While the Sayago et al. (2012) study utilizes an experimental design with ANOVA analysis, my study is not designed to manipulate variables in an experimental manner. Rather, I intend to measure variables and analyse them for correlation. With this main exception, my study design is very similar to Figure 1, and is presented in Figure 2:

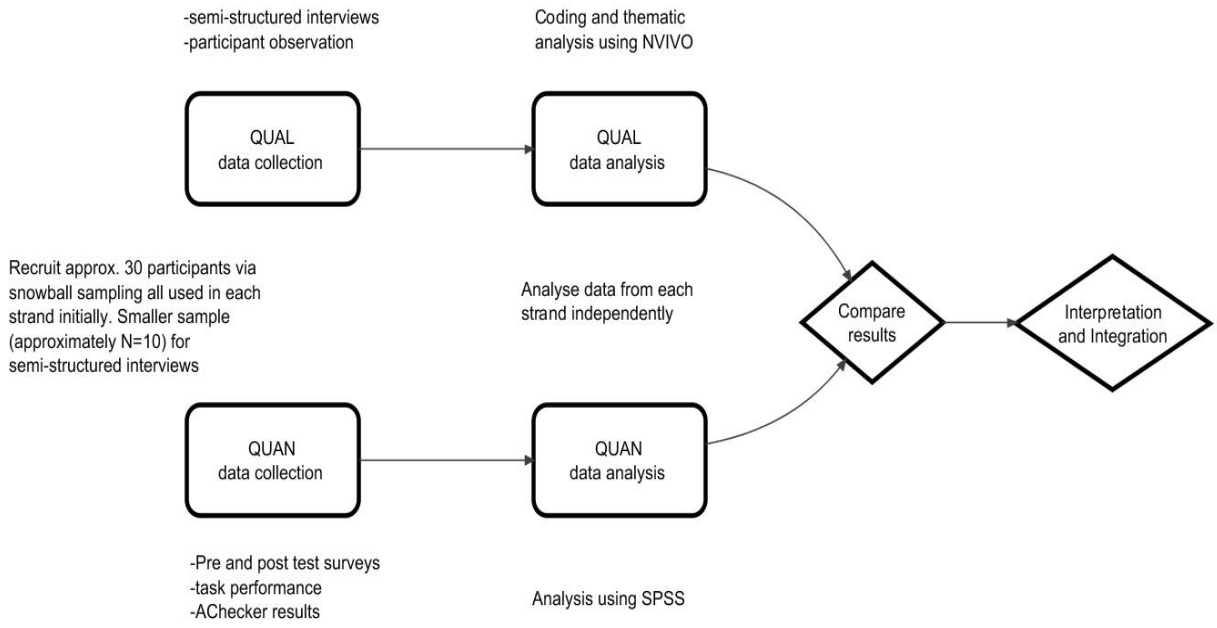

Figure 2. Parallel mixed research design for my study.

Figure 2 represents my study design. In this design, qualitative and quantitative data are collected during and after the performance of tasks by the study participants. Qualitative data is capitalized as QUAL in the collection box, and quantitative data is capitalized as QUAN in the collection box, a notation scheme consistent with their representation in a mixed methods study (Teddlie & Tashakkori, 2009). Figure 2 shows the steps to be followed in sampling through to the interpretations of the results. I have chosen this design in order to explore the research problem by giving voice to the perspectives of participants through interviews, while trying to understand the relationships between the variables that I have identified in the research questions. This multiple approach is required because a singular approach would be insufficient to answer the questions that I have proposed (Creswell & Plano Clark, 2011).

Recruiting

Study participants were recruited from a local community within York Region, Ontario, Canada. Towns and cities within York Region maintain senior centers, and recruiting was conducted through their resources. In order to achieve a large enough sample to provide sufficient quantitative and qualitative data (in view of the technical requirements of my study), I set the criteria for a suitable site to include the following:

1. Minimum active membership of 200 (to provide a sufficient number of possible participants)
2. At least one recent-model computer with high-speed internet access available for the purpose of website accessibility user evaluation
3. Member(s) who can assist in facilitating the selection of a snowball purposive sample (see Collins, Onwuegbuzie, & Jiao, 2007).

A suitable site was located in York Region and consisted of a large physical structure that houses an active seniors centre with a membership of approximately 650 persons age 55 and older. Within the centre seniors engage in numerous activities including playing cards, pool, crafts, luncheons, dancing, darts, bowling etc. A separate room is maintained as a computer centre. This room contains ten desktop computers and a networked printer. All desktops have internet access and shared printing services. Wireless access is provided throughout the centre for those seniors that bring their own mobile devices while attending. The computers used during this study were Lenovo desktops with core2duo E7500 2.93 ghz processors, 2.0 gb RAM, using a 64 bit Windows 7 operating system. The monitors were 22 inch Samsung brand LEDs and a keyboard and a mouse were at each station. The same desktop (machine number G0043) was used throughout the testing.

Pilot Study

My study included a pilot study to determine the usability of the data collection instruments that I used. The pilot was conducted in December 2013. Pilot studies are a necessary step in a research study in order to determine any difficulties in administering and any misunderstandings inherent in the instruments or instructions (Palys, 2003; Rugg & Petre, 2007). For the pilot study I used participants from the active membership of the selected senior centre, selected as a convenience sample (Collins, Onwuegbuzie, & Jiao, 2007) as it is important to use participants that are as similar as possible to the main study participants (Davies, 2007). In determining instrument usability, a pilot study is a formative evaluation similar to a formative usability study (Sauro & Lewis, 2012). The sample size required in a formative usability study depends on the desired accuracy in determining problems with an instrument. Neilsen (2012b) recommends that five participants are generally sufficient at a minimum in a usability study, and I used ten participants in my pilot study.

Ten participants were required due to a number of technical issues that arose during the pilot study and necessitated changes to the use of software tools to better capture data. An initial attempt was made to make use of a remotely configured screen-recording software tool to capture the video of the movements made by a participant on a website. The upload speed of the on-site internet service was insufficient for this tool to be used. An attempt was then made to use locally configured screen-recording software tool to capture this video. This tool became a distraction to the participants, reducing screen size and necessitating my involvement to start and stop the recording for each test. It was abandoned, and participation navigation was captured through observation and written researcher notes and through the internet history recorded by the Google Chrome internet browser.

Participants were first asked to complete a pre-task questionnaire (Appendix A) that gathered background data including age, gender, self-reported attributes (vision, memory etc.), and computer skill level. They were then asked to perform information-seeking tasks on four separate Ontario government websites (Appendix B). After performing each task, the participants answered the positively-worded SUS questionnaire (Sauro & Lewis, 2011, 2012), an adaptation of the original (Brooke, 1996). I told the participants that the session would be audio-taped (consent received) and that any comments they had before, during, or after were most welcome.

The semi-structured interview questions (Appendix D) were designed to expand upon the data acquired from and comments given by participants during the pre-test questioning, task testing, and post-test questioning phases (Appendices A-C). The semi-structured questions were primarily open-ended orienting questions, adaptable to the individual participant, and were not pilot-tested. The lack of a rigid structure permitted flexibility during the semi-structured interview phase and the semi-structured questions in Appendix D served as a general interview guide in the sense that "…topics and issues are specified in advance, in outline form…" (Teddlie & Tassakkori, 2009, p. 229).

One male and nine females participated in the pilot study. The age range was from 60 to 79, with a mean age of 70.60. The former occupations of the participants included teacher, counsellor, engineer, bookkeeper, secretary, real estate, store clerk, and administrative assistant. Computer skills were rated by the participants as Poor: 1, Average: 6, Good: 2, Very Good: 1.

I used one automated software (AChecker, 2015) tool during the pilot study. I chose this tool for its ease of use and open-source access (Gay & Li, 2010). It is commonly used and benchmarked against other tools (for example see Vigo, Brown, & Conway, 2013). AChecker

reports on both known and likely accessibility problems for each webpage tested. It supports testing for compliance with accessibility guideline WCAG 2.0 (W3C WCAG 2.0., 2008). WCAG 2.0 is the guideline used as the standard in Government of Ontario accessibility legislation and regulation (Ontario Regulation 191/11, 2011). AChecker produces a numerical scoring of website errors in three categories: known problems, likely problems, and potential problems.

Table 3 shows those results as a composite score for each website. For example, the participants accessed on average 6.2 pages on the Health website in order to complete the required task (participants used different pages and paths to get to the information or required form). Each of these pages was tested with AChecker, and the total number of known and likely problems is reported here, calculated as adding up the numbers for each webpage. The SUS score is the rating given by the participant for the website, using the System Usability Scale (Brooke, 1996). As an example, on the consumer website, participants averaged 3.6 pages to complete the task. The total known problems averaged 31.80, likely problems 5.6, and the average SUS score was 74.5 (scored out of 100). Table 3 shows the results from this testing:

Table 3.

AChecker and SUS score for websites tested during the Pilot Study.

Ministry	**Pages**	**AChecker Known**	**AChecker Likely**	**SUS Score**
Health	6.20	26.20	.70	74.00
Transport	4.90	31.90	11.20	70.00

Labour	8.70	150.80	1.60	44.00
Consumer	3.60	31.80	5.60	74.50

The number of pages accessed overall (derived from column 2), compared with the SUS average overall scores (derived from column 5) showed a strong negative correlation between the two. The usability rating assigned by the participants as an averaged total for the four websites was inversely related to the number of pages accessed – the more pages used, the lower the SUS ratings, r_s = -.921, 95% BCa CI [-1.000, -.700], p = .000. Likewise, AChecker known problems was inversely related to SUS scores, r_s = -.830, 95% BCa CI [-1.000, -.159], p = .003.This latter is an expected result, given that an increase in the number of pages used increased the total number of errors reported by AChecker on those pages. The total number of pages was strongly positively correlated with the AChecker known problems score, r_s = .939, 95% BCa CI [.763, 1.000], p = .000. It should be noted that these comparisons involve the total overall pages used by participants and the averaged SUS scores of participants for all four websites (composite values). The number of pages used on an individual website and SUS score for that particular website did not always result in a significant correlation.

Another result that might have been anticipated beforehand was the inverse relation between computer skills and time taken to complete tasks. This data was captured by recording the time in my notes during the testing and cross-referencing it with the audio recording time stamps. As individual computer skills increased, the overall time taken decreased, r_s = -.692, 95% BCa CI [-.940, -.116], p = .027. From the demographic data collected, vision has a strongly positive and significant correlation to the participants' self-reported quality of memory, r_s = .783,

95% BCa CI [.153, .997], p = .007 and a negative correlation to the number of pages accessed, r_s = -.643, 95% BCa CI [-.902, -.213], p = .045. Memory was negatively correlated to the number of pages accessed but the correlation is not statistically significant for this sample, r_s = -.592, 95% BCa CI [-.892, -.110], p = .072.

Comments made by participants were recorded during the performance of the tasks and later transcribed. These comments give insight into the thinking of the participants. The theme of adaptation was observed in the qualitative data, through comments made by the participants. This theme involved the adaptation of the user to the website or webpage as presented to them. For example, one participant remarked that, “I know it’s there. When I was doing it, I hit all the colour bits first, then I reverted to the left hand side to the indexes, you found yourself doing it”. Other participants remarked that, “I just kinda go by rote, I’m not thinking while I’m doing it. Ok, I’m getting used to it now”; and “I think after you did the first one, you would find it more easy”. This theme suggests that familiarity with particular websites or particular styles within websites simplified and enhanced the user experience. As reported by the participants, some measure of familiarity was gained during the study. This observation led to the decision to alternate the order of the tasks (see further discussion in chapter four on alternate task order regimens), in order to mediate the learning/familiarity effect to reduce bias in the website satisfaction ratings.

Another obvious theme in the data was that of assistance. The participants in general commented during the study that they sought assistance from family members when navigating websites on their own. One participant reported, “When you have your grandson there, it`s like do this for grandma, so he goes in, does most of that, but I did use the computer when I went

back to university". This was a common occurrence, reported by participants particularly when they ran into difficulties with websites:

> I went into the health one and I couldn`t find anything. I told my daughter in law and she said you have to click here. She is very good with computers. If I have a problem, I just go to her.

The behaviour of seeking assistance with internet and computer tasks was also observed during the study. As one participant remarked to the researcher, "I keep going to ask you which one. I can't ask you anything, eh"?

Conclusions from the pilot study.

This pilot study tested websites for conformance and compared the results with the reported user experience of seniors, while they performed information-seeking tasks on those websites. The findings from this pilot include qualitative themes and quantitative data that support the idea that seniors continue to use internet websites for a number of purposes, including searching for information and performing functions such as banking, communication, and writing. While physical limitations may reduce the user experience, website limitations also have an impact on the user experience. I made a small number of changes to improve the study tools for use in the main study, including alternation of the task order and the inclusion of the use of a second automated software tool in the main study. I decided to alternate the task order for the purpose of reducing any potential bias that might affect the results (Field, 2013; Sauro & Lewis, 2012). The second software tool used was WAVE (WebAim, 2015). This tool was chosen for its reporting of contrast errors. It was noted during the pilot that AChecker was reporting potential contrast errors on the websites. WAVE isolates contrast errors separately, and reports

them in a visual manner as opposed to the more textually based AChecker reporting (see Vigo et al., 2013), making it easier to separate these errors out from other potential ones.

Main Study Strands

Quantitative strand: sampling, data collection strategy, data analysis.

Once a suitable site was located, a snowball purposive sample of at least thirty seniors (N=30) was pursued. I anticipated that this size of sample would provide sufficient data for analysis, given that a snowball sample by definition yields information rich cases and key informants (Patton, 2002). The site selection and sampling resulted in a case study of thirty-four seniors from a particular seniors centre in Ontario. The participants included twenty-six females and eight males. A case approach has the potential to produce valuable analytical conclusions. Analytical conclusions can lead to analytical generalizations of theoretical propositions (Yin, 2009).

From a statistical perspective related to the significance of findings, the sample size of thirty or greater is sometimes considered to be sufficient given the findings of the *central limit theorem* (Urdan, 2010). However, it is also argued that this rule of thumb sizing is not always appropriate, and, in numerous usability studies, a much smaller sample size is sufficient (Sauro & Lewis, 2012). I chose the sample size thirty due to its potential for producing practically significant quantitative results along with its potential for producing rich and deep qualitative data from observations made of, and comments made by, the participants.

Studies of website accessibility involving seniors as participants use the term *ecological validity* to describe experiments that are conducted to be as close to real-world experience as possible (see Sharit, Hernandez, Nair, Kuhn, & Czaja, 2011). I had the participants use a computer within an environment that is commonly used by, and accessible to, participants, rather

than requiring participants to attend an unfamiliar location. As noted earlier, testing was done at a senior centre with computer access, dismissing any requirement to have seniors attend an unfamiliar location using unfamiliar computers.

A pre-task questionnaire (Appendix A) gathered background data from the participants, including age, gender, and computer skill level including internet expertise (Rubin & Chisnell, 2008). Participants were asked to perform four separate information-seeking tasks on four Ontario government websites (see Appendix B). After performing each task, the participants answered the positively-worded System Usability Scale (SUS) questionnaire (Sauro & Lewis, 2011, 2012), an adaptation of the original (Brooke, 1996). This questionnaire provided a subjective measure of usability in reference to each of the government websites in which the tasks were performed. The SUS questionnaire contains ten items using a five point Likert scale. After coding, the SUS produces scores ranging from 0 to 100 in 2.5-point increments (Sauro & Lewis, 2012). I chose the SUS questionnaire because of its reported validity and reliability (Sauro & Lewis, 2012) for the concepts being measured.

As noted earlier, AChecker is an automated software tool that tests web pages for conformance with accessibility standards (AChecker, 2015). AChecker produces a numerical scoring of website errors in three categories: known problems, likely problems, and potential problems. AChecker (2015b) defines its three reported problem categories in the following manner: a) known problems: “these are problems that have been identified with certainty as accessibility barriers. You must modify your page to fix these problems;” b) likely problems: “these are problems that have been identified as probable barriers, but require a human to make a decision. You will likely need to modify your page to fix these problems;” and c) potential problems: “these are problems that AChecker cannot identify, that require a human decision.

You may have to modify your page for these problems, but in many cases you will just need to confirm that the problem described is not present." The reporting by AChecker is in text form as a list of problems under the relevant compliance guideline heading. I used this tool to measure the reported technical accessibility rating of the websites and web pages attended by participants.

WAVE is also an automated software tool that tests webpages for accessibility compliance (WebAim, 2015). It reports on this compliance is a somewhat different manner than AChecker, using a more visual fashion and coloured-icons displayed on the examined page to highlight found errors. WAVE reports contrast errors as a separate category, as opposed to AChecker which reports contrast errors with other errors found, making it slightly more difficult to isolate the contrast errors from the rest. Contrast errors are reported when the contrast ratio is insufficient to render the text and images on a webpage readily distinguishable from the background (W3C WCAG 2.0. (2008). I used this tool to measure the reported technical accessibility rating of the websites and webpages, and in particular for its reporting feature on contrast errors. These automated software tools acted as a post-test heuristic evaluation of the accessibility status of the websites and webpages used in this study.

Table 4 summarizes the four instruments used in the quantitative data collection strand:

Table 4.

List of Quantitative Instruments used.

Instrument	**Measures**	**Reason for Use**
Demographic Survey	Age, gender, education, computer skills, occupation.	Provides basic demographic data for comparisons.

SUS - System Usability Scale (Brooke, 1996; Sauro & Lewis, 2012)	Subjective assessment of the usability of the website. 10 items, 5 point Likert scale.	Provides subjective measure of usability after completion of tasks. Valid and reliable instrument.
Accessibility Checker (AChecker, 2015a)	Website accessibility errors in three categories: known, likely, and potential problems.	Provides an objective measure of website conformance to WCAG 2.0 guidelines.
WAVE (WebAim, 2015)	Website accessibility errors and contrast errors.	Provides an objective measure of website conformance to WCAG 2.0 guidelines.

Data analysis consisted of a number of steps within this strand. The scale results from the SUS were compared to the AChecker and WAVE scores for each of the four websites, commencing with the scores for the home page of the website. I used the error counts for known problems reported by AChecker, and the error counts reported by WAVE in calculating scores. These scores were compared to the individual website SUS scores to determine if a relationship exists. These comparisons aimed to understand whether there are relationships between the reported technical accessibility rating of the government websites and an element of the usability experience (usability) of the participants. Quantitative data were analyzed with SPSS software. Descriptive statistics were developed for initial measures of AChecker error results and questionnaire results. Inferential statistical tools (e.g. correlations) were accessed for comparison of SUS scores to other scaled scores (e.g. time, self-reported attributes) and comparisons and descriptions of the collected quantitative data were made, as discussed in the following chapter.

Qualitative strand: sampling, data collection strategy, data analysis.

Qualitative data were collected through participant observations made during the testing of the participants from the snowball sample (N=34) on assigned tasks. The data were collected in the form of researcher field notes, along with audio recordings of the participants' website usage captured while participants performed the information-seeking tasks on the websites. A coding scheme was developed during the pilot study, and used by the researcher to code observations during the task performance. The scheme consisted of abbreviations used by me to help record places in the individual testing when the participants made comments, accessed a new page, used the back browser, and other functions such as scrolling through the webpage. I also recorded when the participants indicated they had completed the task, or were unable to find what they were looking for. Additionally, each participant was asked post-test interview questions (see Appendix C).

During the data collection with the snowball sample, a convenience sub-sample was drawn (six participants from the main study and three participants from the pilot study) for the purposes of semi-structured interviews designed to elicit narratives from the participants in this sample. The pilot study participants chosen were interviewed along with the main study participants as part of this sub-sample. A convenience sample is defined as "Choosing settings, groups, and/or individuals that are conveniently available and willing to participate in the study" (Collins, Onwuegbuzie, & Jiao, 2007, p. 272). In this study, I selected cases based on the participants' willingness to provide verbal comments, my perception of their interest in the study, their availability, and their expressed desire to participant further. These cases were selected for the purpose of having multiple varied cases from which qualitative data were drawn through the use of a semi-structured interview protocol, designed to elicit rich and deep

narratives of the experiences of the participants. The size of this latter sample was dependant on availability and saturation issues, and in the end resulted in a sample size of nine (N=9) participants.

The semi-structured interviews contained eleven open-ended questions (Appendix D), recorded with a Sony ICD-UX200 digital recorder, and transcribed and analysed with the use of NVivo 10 computer software. NVivo computer software assists a researcher in an open coding process of interpreting narrative data and aligning it into thematic categories arising from the researcher's examination of it (Bazeley, 2007; Kvale, 2007; Richards, 2005). The interview questions were intended to elicit narrative from the participants concerning their website usage. Table 5 summarizes the instruments to be used in the qualitative strand:

Table 5.

List of Qualitative Instruments used

Instruments	**Measures**	**Reasons for Use**
Post-test and Semi-Structured Interview Questions (see Appendices C and D)	Requires researcher coding analysis of answers given by participants.	Provides qualitative data in narrative form for use in reporting depth of participant experience and for quantitizing for comparison purposes.
Researcher field notes and code book	Requires researcher coding analysis of observations recorded during participant task performance.	Provides participant observation data in the form of coded observations and narrative.

The video and audio recordings and the field notes were analyzed by coding with the assistance of NVivo 10 computer software. Coding provided the ability to group observations into thematic categories towards the purposes of interpreting and quantifying this data. The data analysis included a search for themes within the data with the use of open coding initially. Open coding is a process of placing data into categories, indexing it for the purpose of later retrieval (Bazeley, 2007). I reviewed the qualitative data and placed sections of the text into categories based on my interpretation of what the participants were saying in general about a particular event in the testing. I also initially coded by using the language of the participants directly to produce initial categories. For example, if a participant commented that the websites were not easy to use, this comment would be coded under the category 'websites not easy to use'. The categories that emerged are discussed in the findings chapter. Thematic coding joins the categories into central themes. In the software data management program NVivo 10, categories are initially stored as child nodes and child nodes are joined to form parent nodes during analysis. It is through these types of coding processes that themes emerge and become apparent to the researcher during analysis (Bazeley, 2007; Creswell, 2009).

Quantification (data conversion) of the qualitative data was considered in the original proposal for the study. It was originally proposed to count themes and convert themes to numerical variables (e.g. Creswell & Plano Clark, 2011, Teddlie & Tashakkori, 2009) where appropriate, after data collection was completed and initial analyses conducted. Data conversion allows for the use of descriptive and/or inferential statistics to assist in the analysis of the merged data (Driscoll, Appiah-Yeboah, Salib, & Rupert, 2007). The use of statistics permits numerical comparisons of the relationships between variables. I thought that this quantification process might provide insight into the relationship between usability experience (a subjective variable

that was collected primarily through qualitative measures) and technical accessibility measures (an objective variable collected through quantitative measures). However, during the analysis I determined that data conversion was not useful as the counting of themes and words did not provide sufficient data for comparison purposes. The frequencies calculated from word counts, and the finding of six thematic groups in the qualitative data, did not provide for meaningful numerical comparisons, and data conversion was abandoned. I was aware that while data conversion might possibly be useful for descriptive comparative purposes related to frequencies, the qualitative data would be more useful for providing rich descriptions of the perspectives and experiences of the individual participants (Denzin & Lincoln, 2008).

Interpretation and Integration Phases

Each analysis strand produced separate inferences as results. Through comparison, further interpretations and integration of the results were possible. A mixed methods research study requires the integration of findings from both strands, while acknowledging that integration may not mean that a single understanding is achieved (Teddlie & Tashakkori, 2009). It is possible that the results in different strands yield different conclusions, and integration in this situation may require a closer review of the findings or the creation of "…a more advanced theoretical explanation to account for multiple explanations" (Teddlie & Tashakkori, 2009, p. 306). It is in this interpretation and integration phase of a mixed methods study that the quality of inferences is most debated, given the perceived difficulties in integrating qualitative and quantitative methods (Teddlie & Tashakkori, 2009).

An example of the difficulties in integration is given in the study by Sayago et al, (2012). In this study the authors report on findings from the separate QUAN and QUAL strands. The

study reports a statistically significant effect of the ways of marking required fields on the number of errors made by participants (quantitative strand). This finding is based on a statistical analysis of the differences between errors made by different groups using different marking methods. The major finding was that the method of binary classification results in fewer errors than any of the other three methods. The study also reports that interview results (qualitative strand) showed that giving personal data (in required fields) was considered a too personally disclosing activity, and consequently the participants refused to fill out required fields on purpose (refusal effect). It is not reported whether the refusal effect is common and widespread across all four methods of marking required fields or found only in the three methods that don`t use the binary classification. Since the error rates are quite low in the binary classification method, it would seem that an argument could be made that there was less refusal effect in this method than the others. The study has interpreted the lower error rate to be the result of a strategy of purposefully diminishing the cognitive load (through form design) on the study participants. I wonder if there are other plausible explanations, for example, is there an affective component to the binary classification method that in some manner works to overcome the refusal effect?

In my study, quantitative data were analyzed with the assistance of SPSS software. Descriptive statistics were employed for measures of AChecker and WAVE error results, questionnaire results, and task results (completion/non-completion) in relation to research question one and sub questions a) and b). I originally planned to construct a user experience (UX) composite measure calculated from the task results (completion/non-completion), SUS questionnaire results, time, and qualitative data from the post-test interview questions, in order to make comparisons between UX and AC scores. This comparison was to be used to help address

research question two. The calculation of the UX composite measure was planned to utilize the results obtained in the qualitative collection and analysis, and present a context-specific method for this type of calculation. Other studies have attempted to refine composite metrics or indexes (for example see Kim & Han, 2007). Research on UX uses multiple, emergent, and context-specific means to arrive at calculations for UX (Bargas-Avila & Hornbaek, 2011), and I proposed to use the same means, refining the calculation method during the analysis of and the emergent findings from the data. Other recent research (e.g. Brajnik & Giachin, 2014) has further emphasized the idea of context-specific measurables, obviating the desire for a generalized index. These arguments counter the argument in favour of standardized indexes, and I find myself in agreement with it. Consequently I decided to abandon an attempt at calculating a UX composite index.

I did use three of the primary elements of the concept of usability in my comparisons. These elements (Rubin & Chisnell, 2008) are *efficiency* (measured as the time taken on a specific task), *effectiveness* (calculated as completion/non completion of the task), and *satisfaction* (user scale rating of a website). These elements are often used in measuring usability (for examples and discussion see Frakjaer, Hertzu & Kornbaek, 2000; Sauro & Lewis, 2012). As noted, these elements were examined through a qualitative analysis of interview data along with a quantitative analysis of SUS questionnaire data and observations as to task performance. For the SUS I used the positively-worded SUS questionnaire (Sauro & Lewis, 2011, 2012), an adaptation of the original (Brooke, 1996).The positively-worded SUS questionnaire produces scores similar to the original while reducing both respondent and researcher error (Sauro & Lewis, 2012). Researcher error is reduced by removing the possibility of coding errors prevalent in the original SUS (Sauro & Lewis, 2011).

Table 6 summarizes the sampling, data collection, and analyses in relation to the research questions and sub questions:

Table 6.

Summary of Research Questions, Data Collection, and Analyses Procedures

Research Question	**Data Sources**	**Analysis**
1. What is the user experience of seniors when using government websites tested for compliance with accessibility standards?	AChecker and WAVE error results. Post-test questions, interviews, questionnaires, and participant observations from snowball (N=34) and opportunistic (N=9) samples.	Thematic analysis with NVivo 10. SPSS (descriptive). (SUS questionnaire, task completion score, time, qualitative data).
a) What barriers, if any, do seniors experience when using government websites?	Interviews and participant observations from snowball (N=34) and opportunistic (N=9) samples.	Thematic analysis with NVivo 10.
b) How do seniors adapt to the environment presented by government websites while using them?	Interviews and participant observations from snowball (N=34) and opportunistic (N=9) samples.	Thematic analysis with NVivo 10.
2. What are the relationships between the user experience of seniors, barriers experienced, adaptations made, and the reported accessibility rating (of the government websites)?	From snowball (N=34) sample: Usability measures. AChecker and WAVE error results.	SPSS correlation tests. (Variables include barriers experienced, adaptations made).

Limitations

My study is limited in its generalizability due to the smallish sample size (N=34) and sampling method (snowball), from which the quantitative data were collected and subsequent statistical relationships determined. Generally, probability sampling requires larger samples randomly drawn (Teddlie & Tashakkori, 2009), in order to generalize to a population. However, it is most important that a sample be representative of the population about which one wants to make statements (Sauro & Lewis, 2012). In this study, I wish to make statements about seniors that use technology and websites and consequently a purposive sample was chosen, as it is thought to yield key informants (Patton, 2002).

Ethics Review and Timetable

My study involves human participants in a minimal risk setting. The informed consent forms are attached as Appendices E and F, with one form for the pilot and main studies and a separate form for the semi-structured interview, due to the different requirements of each. The timetable for the study is attached as Appendix G. The York University graduate student human participants research protocol (HRPC) was submitted and the study received approval in the fall of 2013.

Summary of Design and Methods

Two separate strands (QUAN and QUAL) of data collection and analysis were used in my study. The main sampling schemes for each strand are different. The larger snowball sample (N=34) provided for the collection of QUAN data. QUAL data through participant observations (live and video-recorded) will also be obtained from this sample. The smaller opportunistic sample (N=9) provided for the collection of participant narrative through the instrument of a

semi-structured interview. The strands were conducted in parallel, resulting in a parallel mixed research design. Analysis of both strands was conducted with the assistance of computer software (NVivo 10 and SPSS). Inferences drawn from the two strands were compared and assessed using an integrative framework (Teddlie & Tashakkori, 2009).

Chapter 4 – Findings

Introduction

In this chapter I detail the findings from the study. The reporting begins with a description of the participants involved in the main study (N=34), moving to a description of the participants selected for supplementary interviews (N=9). I then proceed with the results of the quantitative analysis, followed by that of the qualitative analysis and an analysis of the themes I interpreted from the data. The findings from the two streams are integrated in the discussion presented in chapter five.

Main Study Participants

Thirty-four individuals participated in the main study. The participants included twenty-six females and eight males. All thirty-four participants provided demographic data and performed monitored information-seeking tasks on four different Ontario government websites, with the measures used detailed in Appendices A to C, and the SUS questionnaire (Sauro & Lewis, 2011, 2012). Thirty of the main study participants reported themselves as retired. The other four reported having or seeking part-time work in clerical or administrative occupations. Previous occupations (while working) reported included: clerical worker, government services worker, accounting clerk, teacher, nurse, computer technician, sales, banking, call centre worker, administrative assistant, and town planner. The mean age of the participants was 71 years, with a range in ages from 57 to 87. Normality tests (Kolmogorov-Smirnov and Shapiro-Wilk) and Q-Q plots determined that the age variable was normally distributed within this age range. Twenty-five participants (73.5%) were aged 65 (the normal retirement age) and over.

The majority of the participants (61.8%) reported having attained a high school education level with the remainder (38.2%) having attained a post-secondary education level:

Table 7.

Education level distribution of the main study participants.

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	High School	21	61.8	61.8	61.8
	Community College	6	17.6	17.6	79.4
	University	4	11.8	11.8	91.2
	Graduate School	3	8.8	8.8	100.0
	Total	34	100.0	100.0	

Participants reported using computers for emails, research, bill payment, entertainment, other banking, purchases, games, filing taxes, Facebook, and other social networking/media purposes. Three participants specifically noted that they used the computer for ‘pleasure’, without describing what that might entail. The frequency of computer usage was high, with 85% of the participants indicating they used a computer every day, as shown in this table:

Table 8.

Computer usage distribution of the main study participants.

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Seldom or never	1	2.9	2.9	2.9
	1-4 times a week	4	11.8	11.8	14.7
	Every day	29	85.3	85.3	100.0
	Total	34	100.0	100.0	

The reported computer skills were distributed normally across the main study participants, with a mean skill level rated at average by the participants:

Table 9.

Computer skills distribution of the main study participants.

		Frequency	Percent	Cumulative Percent
Valid	Very Poor	1	2.9	2.9
	Poor	6	17.6	20.6
	Average	16	47.1	67.6
	Good	7	20.6	88.2
	Very Good	4	11.8	100.0
	Total	34	100.0	

Five other physical attributes were self-reported on a seven point Likert-type scale, with 7 being the highest rating (e.g., excellent vision) and 1 being the lowest rating (e.g., very poor vision). As seen in the table below, the means of each attribute were at the 5 and higher level:

Table 10.

Physical attributes of participants as self-reported on seven point Likert scale.

		Vision	Hearing	Memory	Movement	PainFree
	Valid	34	34	33	34	34
	Missing	0	0	1	0	0
Mean		5.62	5.76	5.24	5.76	5.62

Supplementary Interview Participants

Nine participants in total were selected from the pilot (3) and the main study (6) for the purpose of supplementary interviews. The interviews followed the guide found in Appendix D, and I have characterized them as semi-structured interviews. The selection was based on the availability of the participant, the quality (detailed description, helpfulness, interest) of responses made during their initial participation, and a determination made regarding the possibility of the participant being able to provide further qualitative data useful to addressing the research questions.

The availability and interest expressed resulted in three of the selections being made from the pilot study participants, and the other six from the main study participants. The group included one male and eight females, with a mean age of 69 and a range from age 62 to 82. Six participants had achieved high school education, and three had achieved university. Computer use was distributed similarly to the main study participants, while the group reported either an average (N=6) or good (N=3) computer skills level. As with the rest of the participants in the main study group, attributes means (vision, hearing, etc.) were calculated at level five or higher, as participants reported their abilities as good.

Quantitative Results

This section reports on the quantitative findings of the study. Quantitative data were obtained through the use of a survey tool, Google Chrome browser internet history results, time-stamped digital audio recordings, and the use of automated accessibility checkers (AChecker, 2015; WebAIM, 2015). Although primarily used as a qualitative data-gathering instrument, my researcher notes helped to verify quantitative data such as the time taken on the tasks and

whether the task was completed or not. I recorded the times in my notes and compared them with the Google Chrome history results and the audio recordings in order to ensure that I had accurate measures of the time taken on each task and the completion status. The notes were highly valuable given my inability to video record the webpage navigation due to technical difficulty at the site, as discussed previously.

The quantitative data were analyzed with the assistance of SPSS 22 software. The findings incorporate usability study concepts and the measures of efficiency (measured as the time taken during the task), effectiveness (calculated as task completion/rate), satisfaction (user scale rating of the website), and accessibility (using the reported errors from the automated checkers). The individual tasks and government ministries are denoted as Health, Transportation, Labour, and Consumer Affairs.

Efficiency.

Efficiency is frequently indicated as a time measure (Rubin & Chisnell, 2008). I have chosen the time measure as an indicator of efficiency in my study. Table 11 below sets out the mean, minimum, and maximum times (in seconds) taken by participants. These times are broken down by website task (participants were asked to perform four information-seeking tasks on four Ontario government websites listed in Appendix B) and further broken down by incomplete and complete task performance results. In general, a participant spent less time on a task when they were able to complete it successfully, as might be expected. The exception to this observation is the task involving the Labour Ministry. The mean time taken for both incomplete and complete results is approximately the same for the Labour website.

Problems were experienced with the Ministry of Transportation site during the study in the months of March and April, 2014. In six cases the websites pages would not load or were unbearably slow in loading. These cases have been removed for the calculation of composite variables and statistics specific to the Transportation site, leaving that sample at N=28. For all other websites, the sample used remains the same (N=34).

Table 11.

Number of seconds per task based on completion/incompletion status.

	Mean	Minimum	Maximum
Health Incomplete	206	45	570
Health Complete	164	50	460
Transport Incomplete (n=28)	286	120	615
Transport Complete (n=28)	138	40	440
Labour Incomplete	225	50	750
Labour Complete	226	70	450
Consumer Incomplete	209	80	330
Consumer Complete	141	30	285

A related measure of efficiency is the number of pages accessed during an information-seeking task. Table 12 shows the minimum, maximum, mean, and range of the number of pages accessed by participants during their attempts at the four tasks.

Table 12.

Statistics on the number of pages accessed by participants.

	N	Range	Minimum	Maximum	Mean
Health	34	9	3	12	4.94
Transport	28	5	2	7	3.79
Labour	34	10	2	12	4.97
Consumer	34	3	3	6	3.76
Valid N (listwise)	34				

This table shows that the mean number of pages accessed on each website varied from a low of 3.76 (Consumer) to a high of 4.97 (Labour), with a wide range in the number of pages accessed within each site (up to 12 pages on the Health and Labour websites). As might be expected, a significant positive correlation between total time taken and total page numbers accessed was found, r = .434, 95% BCaCI [.061, .713], p = 0.010, with 2000 bootstrapped samples.[1]

Effectiveness.

For an individual participant measure of effectiveness I calculated the task completion rate per participant as the total number of completions divided by the number of website tasks (4). This yielded discrete values of 0.00, 0.25, 0.50, 0.75, and 1.00. The resulting values are shown in Table 13 (N=28). As can be seen in this table, eleven of the participants (39.3%) were able to complete a maximum of two tasks on the websites. Incompletion of a task during the testing was determined either by the participant stating so (e.g. “I can’t find it”), or by my observation that they had not found the required information when they stated they had

[1] Bootstrapping and bias-corrected accelerated confidence interval (BCaCI) are robust statistical measures calculated in SPSS to produce more accurate estimations of both the sampling distribution of the data set, and the 95% confidence interval. See Field (2013) for a full description on pages 198-201.

completed the task. The remaining seventeen participants completed three or more of the tasks. The median value (and mode) for the task completion rate was 0.75, while the median value was 0.6429.

Table 13.

Task completion rates per participant.

		Frequency	Percent	Cumulative Percent
Valid	.00	2	7.1	7.1
	.25	3	10.7	17.9
	.50	6	21.4	39.3
	.75	11	39.3	78.6
	1.00	6	21.4	100.0
	Total	28	100.0	

For a cumulative measure of effectiveness I used the task completion rate calculated for the individual websites (N=28 for the Transportation website, N=34 for the other websites). The task completion rate was calculated as a ratio of the number of completions divided by the number of participants, yielding the values shown in the table below:

Table 14.

Task completion numbers and rates by website.

	Health	Transportation	Labour	Consumer	Totals
Complete	24	22	11	27	84
Incomplete	10	6	23	7	46
Completion Rate	71%	79%	32%	79%	65%

A 70 percent benchmark for successful completion is often used for verification testing of products prior to release, and used to provide information to guide design towards a higher

benchmark such as 95 percent (Rubin & Chisnell, 2008). For this study, only the Labour website failed to achieve a 70 percent completion rate.

Satisfaction.

The positively-worded System Usability Scale (Sauro & Lewis, 2011, 2012), an adaptation of the original (Brooke, 1996), was used to measure the satisfaction level of the participants. Table 15 shows the results from the pilot study. I include them here as support for my decision to alternate the task order during the main study testing. The mean results for the 10 pilot participants are shown for each website. The range of means is from 44 (Labour) to 74.5 (Consumer).

Table 15.

Pilot Study SUS measures.

	N	Mean
Health	10	74.0000
Transport	10	70.0000
Labour	10	44.0000
Consumer	10	74.5000
Valid N (listwise)	10	

This table also shows the original order of tasks (also reported in Appendix B). As noted in the methods chapter, these pilot study findings suggested an alternation of the task order to remove any order bias. Two new task order regimens were devised, based on the pilot study scores. The first task order was from highest SUS mean (Consumer) to lowest (Labour). The second task order was the opposite, lowest SUS mean to highest. These task orders were alternated for the main study participants. Table 16 shows the SUS means for the main study participants.

Table 16.

Main Study SUS measures.

	N	Mean
Health	34	67.9412
Transport	28	67.5000
Labour	34	63.7500
Consumer	34	80.8088
Valid N (listwise)	34	

In Table 16 we see that the SUS means have changed from the Pilot Study scores. The Labour and Consumer websites scores both increased, while the Health and Transport scores decreased. The websites maintained their relative positions overall, with a narrowing of the differences between them, with the exception of the Consumer website which increased its difference from the closest score to it. Notably, Labour maintained its relative position at the lowest end of the scores, and Consumer maintained its relative position at the highest end. Studies of SUS norms have produced a SUS benchmark mean score of 67 for public websites (Sauro & Lewis, 2012).

Accessibility measures.

Accessibility measures were undertaken as a demonstration of the accessibility of the website environments in which the tasks were conducted by the participants. Both automated accessibility checkers WAVE and AChecker produce numerical data on the number of errors located on webpages. Automated checkers provide a quick, easy, and cost-effective way of checking accessibility scores, but can sometimes produce false positives indicating errors where none exist (Vigo & Brajnik, 2011). Automated checkers are in widespread use to determine the existence of accessibility barriers but may benefit by being used in concert with expert human evaluation to determine the full extent of the barriers (Vigo et. al., 2013). In my study, I use the

results from the automated checkers to provide an initial understanding of the barriers reported on the webpages used, in order to provide a context for the digital environment. The findings of the accessibility checking showed that all websites had known accessibility errors within and were not achieving full compliance with the web content accessibility guidelines (WCAG 2.0) level AA (W3C WCAG, 2008).

WAVE data were collected for reported accessibility errors and contrast errors. Those cases (six) where a participant was unable to complete the task on the Transportation website due to a website malfunction have not been removed in this analysis. The purpose of this specific analysis is to report raw accessibility error numbers and their correlations based on the pages accessed by a participant during the task, regardless of whether the task was completed or not.

Table 17 shows the mean number of WAVE reported accessibility errors for all the pages used during a participants' information-seeking task. For example, the number of pages accessed during a Health website task ranged from three to twelve (see Table12 on page 73), and the number of total errors for all the accessed pages ranged from zero to eight. The mean number of errors reported for this task was two. As can be seen, the Consumer website had the lowest mean errors, and the Transport website had the most, over the total number of pages accessed on each website during a task attempt.

Table 17.

Range and Mean of WAVE accessibility errors on all pages used per website

	N	Range	Minimum	Maximum	Mean
Health	34	8	0	8	2.00
Transport	34	29	5	34	8.29
Labour	34	22	0	22	6.56
Consumer	34	4	0	4	1.09
Valid N (listwise)	34				

Contrast errors on the websites show a different pattern. Table 18 shows the mean contrast errors reported by WAVE on all pages accessed per website for each participant. In this case, the Health website had the highest number of contrast errors, and the consumer website had the lowest.

Table 18.

Mean WAVE contrast errors on all pages accessed per website

	N	Range	Minimum	Maximum	Mean
Health	34	60	8	68	21.24
Transport	34	37	1	38	13.94
Labour	34	11	2	13	5.53
Consumer	34	20	1	21	4.44
Valid N (listwise)	34				

AChecker data for known problems was also collected on the websites accessed by the participants during the study. The mean number of known problems reported by AChecker for all pages accessed per website is shown in Table 19.

Table 19.

Mean AChecker known problems on all pages used per website

	N	Range	Minimum	Maximum	Mean
Health	34	46	7	53	19.56
Transport	34	68	4	72	29.97
Labour	34	175	26	201	78.00
Consumer	34	40	10	50	27.79
Valid N (listwise)	34				

The means collected on the three types of errors have been amalgamated into Table 20. This table shows the means for each of the accessibility error category measurements.

Table 20.

Amalgamated error measurement statistics by website.

	HEALTH	**TRANSPORT**	**LABOUR**	**CONSUMER**
N	34	34	34	34
Mean WAVE errors	2.00	8.50	6.56	1.09
Mean WAVE Contrast errors	21.24	12.96	5.53	4.44
Mean AChecker known problems	19.56	29.97	78.00	27.79

Significant and strong correlations were found between two primary accessibility variables measured in the study at the sample size of N=34. The bias corrected and accelerated confidence intervals (BCaCI) were calculated by SPSS at 95% with 2000 (bootstrapped) samples, as shown in the following table:

Table 21.

Correlation statistics between AChecker known problems and WAVE errors.

	r	**BCaCI**	***p***
Health	.526	.224, .832	.001
Transport	.346	.054, .705	.045
Labour	.852	.761, .933	.000
Consumer	.471	.171, .716	.005

These statistics show statistically significant positive correlations between the two automated accessibility checkers on their primary measures of accessibility errors, inferring that they are producing similar findings on these accessibility errors. These product moment (*r*) effect sizes are considered to range from medium (*r*=.346) to large *r*=.852 in Cohen (1992).

Other correlations.

For the statistics related to completion rates, I have removed the six cases where the Transport website malfunctioned. For all other statistics reported here, all cases have been retained (N=34). The bias corrected and accelerated confidence intervals (BCaCI) were calculated by SPSS at 95% with 2000 (bootstrapped) samples (results shown in square brackets) shown in the following table:

Table 22.

Other correlation measures in the main study.

	Completion rate	**Search rate**	**Time**	**Age**	**SUS**
Completion rate	1	*r* = .457 [.128, .700] *p* = .015*	*r* = -.274 [-.516,.022] *p* = .154	*r* = -.388 [-.658,-.047] *p* = .041*	*r* = .036 [-.312,.409] *p* = .857
Search rate	N=28	1	*r* = -.400 [-.844,-.109] *p* = .019*	-.288 [-.588,.009] *p* = .098	*r* = .256 [-.055,.544] *p* = .143
Time	N=28	N=34	1	*r* = .300 [.098,.581] *p* = .084	*r* = -.303 [-.590,-.012] *p* =.081
Age	N=28	N=34	N=34	1	*r* = .043 [-.350,.379] *p* = .810
SUS	N=28	N=34	N=34	N=34	1

Table 22 shows the significant correlations with asterisks above the diagonal in the table. Non-significant findings are without asterisks. Below the diagonal is the sample size used in the calculations. The findings show that completion rate had a significant positive correlation with search rate. Search rate in turn had a significant negative correlation with time. Search rate is a

variable I calculated from observations of how often a participant used a websites' search capacity rather than scrolling through and reading webpages. Participants who always used the search capacity had a search rate of 100% (coded as the number 1) and those who never used the search capacity had a search rate of 0% (coded as the number 0), with gradations in between for search usage of one, two, or three times. Age had a significant negative correlation with completion rate – as the participant age increased, the completion rate decreased. All other relationships were not statistically significant.

Qualitative Findings

Qualitative data were collected via three avenues during the main study. Participant observations were made during the testing of the websites with the assigned tasks, and recorded manually in researcher field note form. These observations included comments, task start and finish times, names of webpages navigated, and whether the tasks were completed or not. These observations were recorded in order to cross-reference with the digital audio recordings and the Google Chrome history record in order to ensure accuracy of the time measures and webpages accessed, as well as to draw my attention to places in the audio transcript for comments made.

Digital audio recordings were made with a Sony ICD-UX200 digital recorder during the testing, and further digital audio recordings were made during the semi-structured interviews of the smaller subset of the test participants. This data were transcribed, coded, and analyzed with the assistance of NVivo 10 computer software. Field notes helped with the transcription process, by clarifying the start and end points for each performed task and allowing me to highlight comments made by the participant for ease of reference while transcribing. The next sections deal with the findings from the qualitative data, coding, and analysis.

Coding Results

Preliminary coding of the text transcripts produced thirty-six individual code categories, known as 'nodes' in NVivo 10. In the initial coding I grouped pieces of text under nodes that emerged as I read through the transcripts, sometimes using code titles that were directly taken from the text. For example, an initial code "I'm missing something" was assigned when I found it in a transcript as an utterance made by a participant. Other code titles were developed from my sense of what idea(s) the participant(s) were addressing. For example, the initial code "Navigation Issues" captured transcript utterances made such as having difficulty finding desired information, difficulty in scrolling, cluttered webpages. Specific pieces of text were sometimes initially coded at several nodes. For example, a reported navigation issue might also be coded at the initial code "Lost", as well as "Jumping around", and "Navigation Issues". This initial code list is shown in Appendix H.

During secondary analysis of the data, I merged relevant and related nodes reducing the number to twenty-two individual nodes. These nodes were further analyzed and I linked them under six thematic groupings. The six thematic groups are:

1. Adaptation
2. Feelings
3. Technical Issues
4. Navigation
5. Evaluation
6. Suggestions

During the writing process the nodes were further examined and more merging took place, with the number of individual nodes being reduced to nineteen and the number of themes being maintained at six. The next sections examine the contents of each theme. When transcript quotations are used, the abbreviation M refers to a main study participant, and SSI refers to a semi-structured interview participant.

Adaptation

This theme arose from observing participants as they worked through their tasks, the recorded comments made, and the responses they gave to question number nine in the semi-structured interviews. This theme also emerged during the pilot study. I think of it as adaptations made by the participants to the digital environment, noting past and ongoing adjustments. Within this theme are four sub-themes: previous experience, persistence, primary strategies, and using a different interface device.

Previous experience.

The theme of previous experience was found throughout the narratives of the participants in the study. Participants were quick to discuss their history with computers, and quick to articulate what and how they had learned from their previous experiences. I interpreted these articulations as involving an adaptation by participants to digital environments, in this case based on their previous experiences with them. Participants reported on two main drivers of this experience – the workplace use of computers and self-taught use outside of the workplace. For example, SSI7 (female, 61), stated:

> My first experience with computers was, let me see now, about 1976. I started with a word processing program with word perfect - fair size machine, green screen, orange

> type. It was an actual computer, and I just had it at work. I didn't have a computer at home until I guess with the internet and everything maybe it's been 20 years with the home computer.

This participant's experience was typical of those who had worked during their adult lives in occupations including clerical, accounting, or administrative natures. Participants noted the use of now-defunct programs such as word perfect, Lotus 1-2-3, and software programs specific to their employer. In contrast with those who had workplace experience were those who learned computer use at home first and later applied that knowledge upon re-entry into the workplace. SSI5 (female, 68) stated that she used computers in her home long before returning to the workforce:

> I have been at it forever, we've always had one. We had the old clunker Apple when they first came out, in the house. An old, old, old, apple and that's what my husband started the first ancestry on. It was my brother's keeper, was the name of what that was, and you could do ancestry searches through that. That was the first attempt at it, so we had that the little apple, then we went on to the Macintosh, he loved the Apple family, and I liked the other one, so when I went to work I guess probably in 28 years ago, the only thing I worked on was a computer.

Previous experience of a non-computer nature was noted as being of help when searching through websites. One participant felt strongly that the ability to navigate a website was enhanced if one had previous experience in the subject matter of the navigation. M14 (female, 65) described this effect while navigating the Ministry of Labour website:

> It's not easy to find, unless you know to go into the search. Yeah, it really didn't say anything on the side, it had health and safety, which I would know that because I have

> worked, otherwise, like I taught WHMIS [Workplace Hazardous Materials Information System], I would go to the health and safety, but the normal person wouldn't not necessarily know that, I don't think. So it just depends, in general, what you know on where to go to look.

Her conclusion is that knowing what to look for makes the navigation easier, and previous experience in the subject field is of benefit, helping with the understanding of the terminology involved. Similarly, those who used government websites previously felt that it was easier to navigate the websites used in this study. M15 (female, 67) allowed that the first time she had experienced the Ministry of Transportation website she had significant difficulty:

> I just went through this one. It was a pain in the ass. I had to put vehicles in my name, finding the information for what I had to do, I was ready to throw it. And then my sister-in-law phoned and said this is the one, but it's not the one, but we eventually found it.

Being self-taught and learning by trial and error was reported by the participants. These adaptations included what might be thought of as learning through experimentation, wherein learners experiment with the environment and attain solutions through what the environment has to offer and their experiences with it. For example, M17 (female, 66), indicated that there was a time when she would be totally lost during an information search on the internet, but she had learned to use the search facility to narrow the amount of information presented. This adaptation was not obtained easily. As this participant stated, it was arrived at with a measure of frustration and anger:

> Two clicks and there I am, search, right? Well you see I never used to do that, but I have been using internet for a while now, and I don't go to all these things that I used to. It used to drive me mental, because I would be everywhere in the internet except where I

> really wanted to be, and so, this to me is key. I think at one time, I think we've got a little smarter, in our, because if I asked a young person, I'm looking for seniors or something, they'd go right to where it is. I would think, why didn't I do that, and I got angry at myself because you know, they know more than I'll ever know.

M18 (female, 80) notes a different form of adaptation. She moved from the position of giving up on locating information on the internet to accepting it for what it is, and uses what she refers to as the just keep looking approach. As she terms it:

> I'm getting better at looking things up on the web. I would get so frustrated I would go and find a book to get the information rather than trying search on the web, but I'm getting better. I think that probably the site has so many things to choose from, like I say you just have to keep clicking on things until you finally come to what it actually is you are looking for. I bragged to my son last Sunday that I had found two things on there on the internet, I felt quite proud of myself, I did it myself.

During her performance on the study tasks, M18 used the search function on only one task whereas M17 used the search function for each of the four tasks. By comparison, M18 was able to complete all of the tasks, while M17 completed three of the four.

Persistence.

Participants commented on their ability to persist until they located what they were looking for. Comments ranged from "I think I would have eventually got the health form, I do eventually get there with these things (M1, female, 72), to "At home I would just stay with it until I read everything there was…I'll find it, poke around the program" (M10, male, 82). In contrast to these comments on their positive ability to locate information, were those that

indicated they would not persist, that they would quit after a short time and find the sought information elsewhere. For example, M11 (female, 66) stated that:

> I may have got there eventually. I wouldn't waste that much time sitting there thinking about it. I would go to Service Ontario or Service Canada (physical locations), whichever place. I didn't like this site at all. I didn't like that they put so many things that were irrelevant after I clicked on what I wanted… I don't think I would have to learn anything new; I would just have to take a lot of time doing it.

These comments on persistence (and lack thereof) are demonstrative of an adaptation to the website environment, dependent upon the will of the individual involved. Some participants were adamant about their desire and ability to locate what they were looking for. I sensed that these participants were not willing to let the website or system defeat them in their purpose, and they would soldier on to find what they were looking for, no matter how long it took. The other approach to persistence was to adapt by seeking the information through another format, in the face of an inability to find it in what might be considered a reasonable amount of search time. A suggestion from M15 (female, 67) might help to bridge the gap in the understanding of the differences between these two methods:

> I think they could be easier…some of the things spelled out quicker, especially for seniors. They're easy enough to use but I find that a lot of time you have to go through more pages to find what you're looking for, like if you want this one or you want that one, you just have to keep going to find the one you need.

Persistence was seen to be an important element in the participants' strategies and willingness to use the internet. One group of participants reported and were observed to have adapted to website usage and any difficulties encountered by being persistent in their efforts, and refusing to

yield to failures, while another group gave up easily and moved on to other means.

Primary exploring strategies.

I observed two primary exploring strategies used by the participants during their task performance. The first strategy I have labelled the 'reading strategy'. This strategy consists of the participants reading the content of the webpages visited, scrolling down when necessary, while exploring the webpage for the sought information. The second strategy I have labelled the 'search strategy'. In this approach the participant typed words into the search box area on the webpage, and then perused the listed results to find one appropriate to locating the sought information.

In the reading strategy, if the sought information was not found on a webpage, the participants would then click what they considered to be a suitable link leading to another webpage, wherein the reading process would commence again. A participant (SSI7, female, 61) described this process as: "I just keep pressing buttons until I get what I want. I just move the mouse and I click on different things until something looks like whatever it is that I'm looking for". This approach does have its downside in the amount of information that is available on any webpage. As M15 (female, 67) notes: "You see there's all this here, but I find with the government websites there is so much information that you have to go through so many pages to find what you're looking for".

M10 (male, 82) has adopted the strategy of reading everything to find what he is looking for. As he says, "Normally I would try to read everything on the (webpage), every subject that came up for every heading". While this strategy may sometimes yield success, others have gone to the search strategy to reduce the time spent on information searching. As M17 (female, 66) stated, "There was a time when I would be all over the place, I would never know when to get

what I need, I'd be everywhere but the right place". As a result of this sense of 'lostness' (Smith, 1996), M17 has learned to recognize the search function as a tool to overcome the problem. M30 (male, 80) has taken the simple webpage-provided search capacity a step further, acquiring the skill to use a Boolean search. It is his opinion that this search technique further increases his efficiency and effectiveness. As he says:

> I suspect that most people have a bitch of a time finding what they're looking for. I have got exceptional skills in that kind of search, because I have learned the Boolean search, and without that [you will have difficulty finding information], it doesn't matter what you're looking for.

M30 felt that government websites in general were difficult to find information on, and he held forth that his skill level was above that of others, suggesting that the majority of seniors would have great difficulty navigating these websites. The quantitative data somewhat supported his contention, given the task completion rate of 65% over all four websites.

Using a different interface (using a different type of machine).

A small number of participants commented that they had traversed from desktop computers to tablets and/or smartphones, preferring the touch screen ability found on those devices, particularly for the use of games, email and information retrieval. SSI2 (female, 68) felt that her smartphone excelled at information retrieval, and used it as her interface of choice. Not all agreed with this choice however, referring to the perceived limitations of these devices. SSI3 (female, 80) thought that it might be difficult to use a favoured program: "I don't know how you would work with excel with numbers on a touch screen", while several other participants felt the screen size on a smartphone made it difficult for any type of text entry. SSI5 (female, 68), when describing her son's attempt to have her acquire a smartphone, noted her response to him:

> I said it's got such a tiny screen…the thing is for my computer I don't have to put these on all the time (glasses), but that little screen I wouldn't be able to read it I would have to wear my glasses continuously.

SSI9 (female, 65) has adjusted to the small screen size by using a stylus for typing, and demonstrating her use of it to me:

> This one I like because it has the stylus….especially if you're sending a text. If I touch here I actually have a keypad, because I find my fingers are too fat, and I get frustrated, so I use the stylus for almost everything I do. It interprets handwriting too.

While some participants, as noted here, preferred a touch screen device, many preferred the use of traditional desktops and laptops.

Feelings

Participants expressed feelings in and about their use of the internet and websites. These feelings largely appeared in two forms. The forms are described in the next subsections as a sense of frustration with the websites, and alternatively a sense that the participant was somehow at fault for the undesired outcome of the website usage. These two forms can be thought of as outward-focused (frustration with the object) and inward-focused (self-confidence issues).

Frustration with the object.

A sense of frustration with websites during previous usage was reported by participants. This frustration was a result of being unable to complete or achieve what was desired by a participant when using a website, and often ended in the user giving up their attempt. As M18 (female, 80) put it, “When I search the websites I get very frustrated with sites… (once) I got so frustrated…it would let me go so far into it and then it would say couldn't do it, and so eventually I gave up”.

SSI5 (female, 68) described her general experience with computers as occasionally frustrating. She stated that, “Sometimes I get frustrated when I want it to do, like I know what I want it to do but I can't make it do it”. This sense of frustration implied that the computer controlled what she wanted to do, and that she was at its mercy, as in her statement, “If it's a new program for something that I want to download and that, and it's just saying, your server won't allow you to do this, and why won't it allow me, I want to do it”. SSI6 (female, 62) also felt at the mercy of the machine, reporting that the plethora of passwords required to access what she wanted drove her insane. “I never feel confident using any website” said M11 (female, 66). Her utterance reflects on her poor experiences with websites, and her lack of trust in them, as she also referenced previous experiences with websites crashing, as she put it, and leaving her unable to complete her searches.

This sense of being at the machine’s mercy produced feelings of inadequacy, as in M10’s (male, 82) expressed during his attempt to find information on one of the tasks in the study,

> Well, I tell you I feel stupid, downright stupid. I read the question, I don't know what the heck I'm looking at anymore. I try to eliminate, I read a heading, but it's not it. You want my opinion? I feel like a dummy.

This type of feeling is echoed in the next subsection, wherein participants report on frustrations with the self, and feelings of inadequacy towards computers.

Self-confidence.

In describing her inability to locate the desired information on one of the study tasks, M17 (female, 66) denied that the website or computer was at fault. “As far as I'm concerned, I don't think there’s anything wrong, I think it's me. I should have right from the start, when I didn't see what I was looking for” (commenting on using the search bar). She wondered aloud if

she was normal, and if her approach to using the website was normal. As she proceeded through the tasks, and met with success in a latter one, she uttered "I'm beginning to feel smart!" This progression from a feeling of blame to a feeling of success was indicative of the theme of blaming (or praising) the self for the results of website interaction. Generally, the feeling was seen when the result was negative, when the task was not completed. In those circumstances, this type of participant would blame themselves for the result, citing a number of inadequacies including lack of experience, lack of confidence, and lack of interest.

For example, SSI1 (female, 64) assumed that she did not have the skills to perform the study tasks, "I wouldn't have tried it otherwise, I just assumed it was going to be beyond me". And M18, (female 80), while describing her performance, stated: "For me using it, is not too bad at all", while suggesting that her skill level was poor, and that she had performed rather well given her low abilities. M21 (female, 61) described herself as lacking in both confidence and practice: "I just need practice. Well, confidence, you know I'm not that confident". M23 (female, 74) offered the view that, while limited in her skills, she derived great satisfaction (and a measure of humour) in her persistence:

> I know I'm not right up there, but I love it. Sometimes if I don't know something...I'll sit up there and try to figure out something and my husband says why don't you just take it down the road to (the computer store). But I say I'll sit here until I figure it out, it's the great satisfaction of being able to do it. If I could only figure out the sewing machine.

Technical Issues

Two technical issues either arose or were noted during the study. The first issue involved the Ministry of Transportation website, and was an issue that occurred while a few participants were attempting to perform the information seeking task for that website. I have labelled this

issue ‘slow computer’. The second issue was reported by participants as they used the Ministry of Health website, and was related to the colour contrast of items on one of the webpages. These issues are discussed in the following subsections.

Slow computer.

During the study and for an approximately six week period in the months of March and April participants had difficulty with the Ministry of Transportation website. The difficulties included webpages on this particular website that would not load, or were very slow in loading. Six participants experienced webpages that would not load, and were unable to complete the task as a result (these six results have been excluded from statistical analyses). I was unable to ascertain the cause of this, and the only remedial action taken was to slow the pace of the study down in the hopes that the issue would rectify itself. By mid-April the website was functioning well. The difficulties produced a number of comments by the participants involved, and somewhat surprising to me, it was not the first time participants had experienced this. For example, M16 (female, 59) remarked that when in the past she had experienced problems with non-functioning links, she would respond by not returning to that website again. M11 (female, 66) felt that “it’s like when you try to phone them, please hold” and offered that the difficulties with websites was another reason to go directly to a physical office location.

M18 (female, 80) had experienced a similar problem with her home computer while attempting to change her address on a government website, and responded in this fashion:

> I gave up and went to the place in Newmarket, Service Ontario, and she made me do it on the computer, and down there it worked ok, so she blamed it on my computer. But I went home and tried it on my computer, and it still wasn't working, in the site.

These difficulties with the Ministry of Transportation website did not deter the participants from

completing the remainder of the study, and I observed that it appeared to be something many had experienced before. That said, previous difficulties such as these seemed to reduce the confidence that participants had in using government websites. As M11 (female, 66) said, “Just the day you need it, it crashes and then you're stuck - what do I do now?”

Colour contrast on the Health website.

The second major technical issue that occurred during the study involved the Ministry of Health website. One webpage on this site used a combination of colours that caused participants to have difficulty reading it. This webpage contained the link to the much sought OHIP card renewal form. Participants had difficulty locating the proper link on this page (the WAVE contrast report for this page noted multiple very low contrast errors indicating a potential for difficulty in distinguishing text and images from the background). M22 (female, 74) remarked that:

> That was almost easy...if I'd looked up and seen it in green, but I was looking at the list still…I don't think I'd use this frequently, relatively simple except for the colour, easy to use, until I got to that point.

M24 (female, 82) also had difficulty with the colour, “The whole thing, is this the form? It's green too. When I saw that, it should have jumped out at me right? Because that's what I was looking for”. For M19 (female, 68), the page caused her problems in picking out the required information due not only to colour but also due to what might be called congestion: “There's the form. See, I was looking in here, form link, there's so many things, you have to read every little thing, sometimes you just glance over, I guess that's what I did”. As noted in the quantitative findings, the Health website had the overall highest mean contrast errors.

Navigation

Comments and observations on navigation issues were made during the study. The main observation was that of being lost, wherein the participant was observed and/or commented on their having navigated to a point where they weren't able to progress further in their search for information, due to not knowing where to go next, and, in some cases, being unsure as to how to get back to where they had navigated from. The other main entry in this thematic category consists of comments made about the amount of information available on the websites. This section presents the findings on these two items.

Being lost.

M10 (male, 82) felt that he was missing something that should be there, on the webpage, when he reached a point of lostness:

> Well, I don't see anything like that here, unless, what have we got there. You know you read some of this stuff you don't, you kinda figure you're missing something. I'm not as smart as I thought I was. Well there's nothing here about forms, nothing about renewing, services, programs, it wouldn't be under that. Let's try this thing. No, there's nothing there either. Well sir, I'm afraid I'm at a loss.

His feeling was echoed by M11 (female, 66) who found herself on a page with material that wasn't what she was looking for, and remarked "…now I can't figure out what to do". And M13 (female, 76) thought that she may have clicked on a wrong link, and ended up disoriented and lost:

> I must have pressed something. I don't know. I can't see what a worker would have to do. I can see where they can read things, but I can't find my arrow, oh there it is ok (sighing). I guess I'll go there. I don't think I'm getting anywhere.

These three participants, like the others who reported being lost, were unable to complete the required task. The feeling was further described by M27 (female, 62) as “I’m just going in circles”. M32 (female, 57) described the issue as a matter of how the webpage was structured and laid out. In referring positively to the Ministry of Consumer Services website, she noted: “I like it, it’s simpler, laid out nicely. It doesn’t have all the places to get lost”. This website received the most number of positive comments, and notably also, the highest score on the SUS questionnaire.

Too much information.

Numerous comments were received about the amount of information found on the websites and webpages used in the study. For example, again talking about the Consumer Services website, M14 (female, 66) noted:

> It's there if you need to find it, if you can find it, the consumer one was the easiest one to find the information you were looking for. The other ones need to be worked on so that they’re more user friendly. They're not bad, there's lots of information there, it is available.

M15 (female, 67) felt strongly that there was too much information available:

> There's so much information that you would just have to go through and find what you were looking for. The problem is that there's so much involved that you can't really. It’s hard to find a contact that you could even e-mail somebody and say this is what I'm looking for, I'm having a problem, could you tell me where I could look to find it better. I just find with the government websites there is so much information that you have to go through so many pages to find what you're looking for.

While the matter of too much information was primarily a common complaint made by participants, occasionally it produced a ‘give-up’ response. M16 (female, 59) remarked that:

> Oh my god. This is just too much, too much, it's not easy enough. Ok. I should have gone here first….I'd probably just give up and, and I'd just contact them if they have an area to write out a letter, and then I'd just ask them a question like where I can get on the site because it's just way too much reading and it looks like it's more for university educated or whatever.

For M16, not only was there too much information resulting in too much required reading, the quality of it was such that she perceived it to be in a language more attuned to the university educated, as opposed to being in the more common, colloquial, language she preferred. Using a more conversational, straight-forward language would appeal to M16, and make the navigation of these websites a better experience for her. As she summed up, “you know, they’re really not simple enough for ordinary people, it’s too much wording”. She felt that her friends and family would never use these websites because of this issue.

M19 (female, 68) was willing to make allowances for there being too much information on the webpages. She stated that: “I find anything you want to know, you just go there and they tell you. Of course they give you a lot more information than you need, but yeah, it's very informative”. However, as stated earlier, she also felt that the amount of information did cause her to sometimes miss what she was looking for, as there were so many items on the pages and occasionally she just glanced over them.

M26 (female, 81) proposed that websites be devised to target seniors in order to get past the information issue. She felt that:

> I think the wordings could have been a little more clear, for people my age and over,

> trying to peer away trying to figure out where things were, like I said. If they had one just for seniors alone, and then that would capsulize it.

Her thinking was that seniors didn't need all the information presented on the webpages, and could be directed to areas that were more topical to the age group, thereby alleviating the need to search through everything presented in order to find what they were looking for. M27 (female 62) made similar suggestions, thinking that the use of better signposts to distinguish areas of interest to seniors would be useful. For example, in discussing the Ministry of Transportation website, she remarked: "There was nothing there telling me up to 79, there should have been something said…over 80, this is what you do, then I would have found it no problem".

Some participants felt that it was the subject-matter of the websites that forced them to be complexly worded. For example, SSI3 (female, 81) noted: "Some of them have almost too much, too many choices, but I suppose because it is so complex that you have to have a lot". And SSI4 (male, 82) thought that the webpage writers could consider the abilities of their senior readers more:

> I don't think they go out of their way to make it hard, maybe some of the people that write these programs assume that you know what you're doing, a lot of people don't, and they don't read it so as a result they get on there, they make mistakes.

In summary, navigation issues, as reported from the participants' perspectives, existed in two primary forms. The first was the propensity to get lost while navigating, and the second was the matter of too much information on the webpages.

Evaluations

This theme arose from comments made by participants that served as evaluations of websites and the participants' relationship to them. The comments in this theme included such

things as evaluations that the websites in general were either easy or not easy to use. Further exploration of the comments yielded a sub-theme which I interpreted as attempts to convince me of the participants' competence with internet usage. Also noted was a sub-theme based on comments about the motivation to use government websites related to the perceived usability of them. This section deals with these three sub-themes.

Websites are good, easy, or not easy to use.

Comments in this area included outright statements that government websites were good, excellent, easy/not easy to use, and in general contained a lot of good information. For example, M20 (female, 84), while discussing the Transportation website, indicated that she felt there was a lot of information available there, and that she had used it in the past, and described it as being excellent. M25 (female, 70) described the Consumer website as: "This is all great information...I think this website's really good...it's full of information". M34 (female, 73) also described the Consumer website as excellent, while qualifying her evaluation in a somewhat negative manner based on previous experience with government websites: "I have used government websites before, they're not bad, don't drive me crazy". And M10 (male, 82) felt that government websites in general were good:

> Well they're good, sometimes you got to go through two or three of them before you find out what you're looking for, you find one, and then they got a reference to something else, on the average they're as good as anything that's out there, I would say.

M14 (female, 66) distinguished between easy and not easy websites, determining that the Consumer website fit the former evaluation while the other three tested websites fit the latter. She was enthusiastic in her praise of the Consumer website:

> That one was easy to find. You can find the information very quickly on that one. That's the only one, it was in their popular topics drop down menu. You can pick anything you want for which one happens to be affecting you. That's the only one that was actually easy to find. Just scams in general. That's an easy one. A different person probably set it up. This one was an easy one, the first one that was listed in the drop-down menu. Then it gives you a whole variety of which ones that you have out there.

M32 (female, 57) was also supportive of the Consumer website, indicating that the ability to quickly find the required topic through the menu list saved her from having to use the search function. As a result, the described that website as: “Nice and clear, a lot more intuitive than the others, especially for seniors”.

On the other hand, numerous comments were received describing the websites as not easy to use. M16 (female, 59), as referenced earlier, described the websites used in the study as not being easy for ‘ordinary’ people to use, due to the difficult and lengthy wording involved. M24 (female 82) also found the websites difficult, stating: “That wasn't simple, I didn’t think it was easy either. There was so much, too confusing for me”. And M29 (male, 61), in describing the Health website, felt that the requested information should be easier to find.

Motivation and perceived usability.

Closely related to the description of websites as good, easy, or not easy, is the reported usability of them and the motivation to use them. In the words of M1 (female, 72), “Well, I wouldn't want to look up some of these things, on the last couple of sites, not the greatest, that labour one is terrible”. In her estimation, the websites were not easy thereby limiting her motivation to use them. For M14 (female, 66), her motivation to use government websites was directly related to perceived incentives found in doing so:

> I go to the federal ones but not the provincial ones. I have gone into the federal one I guess because they probably give you more money, you get your government cheques, because I also ran a business from home, I got a GST number and I go into there for taxes. So it would be federal rather than provincial.

She remarked that she didn't think the Ontario government websites would be much use to her, thinking that if she were over 80 she wouldn't be using computers anyway, and therefore wouldn't be looking for information such as how to renew her driver's licence at that age.

For M13 (female, 76) the lack of motivation to use computers dates back in time to the 1970s. During that time she developed a distinct dislike for computers, and has carried that lack of motivation to use them forward to the current day. As she puts it:

> I know I'm lacking in anything to do with computers and quite honestly I don't think I'll get any better. I blame it all onto my first husband - in the 1970s he was computer mad, and he tried to put my housekeeping on the computer, my chequebook on the computer, and he used to sit for hours, and say, just look at this, and I sat next to him for hours, and he didn't even know I was sitting there. I blame him for my dislike of computers. Well, yeah, he was so clever, and I just hated it. And I still don't like it.

M13 is not alone is this historical dislike of computers. M12 (female, 62) had a similar lament, noting that during her introduction to computers several years previously, she had developed a dislike for them that she had since overcome to a degree, now using them for emails, games, and Facebook. However, at the onset of her experience she was not at all enamoured of them, finding them too absorbing of attention. As she stated: "I used to watch my husband, and say, 'that's why I don't like computers".

A need to be perceived as competent.

An intriguing insight into some participants' self-evaluation centres on what I think of as an expressed need to be perceived as competent in the use of computers and the internet, and information-seeking on websites in general. Earlier I referenced M30 (male, 80), who enthusiastically discussed the use of a Boolean search method that he used in his search activities. He also took time to comment on his own abilities, reporting that his abilities exceeded that of the senior population in general. Several other participants also tended to separate themselves and their abilities from the rest of the general senior population. M4 (male, 72) remarked: "I guess I must be at one end of your spectrum of users. The other end is what's a website. I'm not sure my wife would be able to finish this test".

M34 (female, 73) positioned herself as better with the computer than her husband: "not everybody is comfortable with the computer...some are more comfortable, I know my husband is not comfortable at all, he has certain things he uses it for and other than that, that's it". She also claimed to be better than the average senior:

> For me it wasn't too bad because I know about the search, but the average senior wouldn't know how to do that, they would be sitting there for hours trying to navigate, I think, this was terrible, I didn't like this one at all.

While her conclusions about having superior skills might well be true, I interpreted her comments and demeanour during the testing as also containing a desire to distinguish herself from the senior population and a wish to appear competent with computers. Early in the study I had been questioned by many participants as to how well they had performed on the information-gathering tasks. I informed them that the testing was of the websites, not of the participants, yet it appeared to me that many participants were more concerned with their own performance (in

terms of such things as successful completion and time taken) than how well the websites performed. And others felt that they were competent in their computer use given their advanced age, and wanted me to know it. For example, M20 (female, 84) went at length to describe her use of both her laptop and her iPad, noting that she travelled with her iPad, using it to complete her banking and send e-mails while away. She summarized her story by claiming, "Not bad for 84, I guess".

Suggestions

This theme arose as the result of examining responses to the final two questions during the semi-structured interviews. The semi-structured interview questions were presented to nine participants selected from the main and pilot studies. The first question asked for suggestions on how to make government websites easier to use, and the second question asked for any other thoughts about seniors using government websites. Both questions produced numerous comments and ruminations on how to improve the websites and the internet in general. The findings from the comments made in response to these questions are presented in this section.

Suggestions for government websites.

Five of the nine semi-structured interview participants offered direct suggestions to improve the ease of use of government websites. The remaining SSI participants either did not make suggestions, claiming the websites were easy to use (three participants) or lamenting the difficulty of using them (one participant). The latter participant (SSI4, male, 82) had this to say:

> I don't know, it seems to me that a lot of their sites are talking around in circles, just get to the point, and do it that way. I mean if I'm looking for a way to tie up my shoelaces, maybe they should be a heading there, how to tie your shoelaces, right over left or left over right, instead of some rigmarole that…now I can't say too much about that.

Of the five with direct suggestions, the general theme was that the websites should be made easier to use, especially for seniors. It was suggested that there be fewer main categories, with subcategories leading from them:

> Well I could say initially there should be fewer categories, and then each of those categories has their own categories, so you get in to where you want to go but in a simpler way I think, you know, instead of having all kinds of categories in the first place, maybe we could have three or four basic, and then each one of those would have three or four and kinda work your way down or up the ladder, whichever way you're going. (SSI3, female, 81)

This idea of simplifying the website was also noted by SSI5 (female, 68), who felt that simplicity meant the reduction in the number of steps required to find the desired information. She claimed that, in the case of forms, once one got to the form it was easy, but the finding of it was difficult, something she attributed to there being too many steps in between. In support of making the websites simpler, SSI7 (female, 61) proposed an alphabetical index for every topic covered, situated on the main page, and with clickable links. She emphasized the use of an alphabetical order, in her words: "I think especially for the sites that older people would use, I think if it was alphabetical then clickable, then it would be where it would take you right to where you want to go". SSI9 (female, 65) thought that more drop-down menus would be useful in finding information, while both SSI5 (female, 68) and SSI6 (female, 62) thought that emphasizing the search function was a good way to improve website usage by seniors.

Thoughts about seniors and government websites.

Each of the semi-structured interview participants offered thoughts about seniors using government websites. Although SSI2 (female, 68) reflected that she did not generally use these

types of websites, it was her opinion that seniors would benefit by using devices such as the smartphone. As she stated,

> I find with the phone it's easier to find websites than on the physical computer. With the microphone, you talk into it and you just say a couple of key phrases and it'll bring up what you want, whereas I find that on here (referring to the desktop), it doesn't bring up the related sites.

SSI3 (female, 81) remarked that: "Some seniors, it's like anything else, some seniors sort of adapt to the computer quite well, some not so much, and some try hard". She was of the opinion that an exacerbating issue with the computers was the occasional slow loading of webpages. It was her observations that this slow loading resulted in a clicking of other links, which would take a person to a webpage that wasn't the desired one. Arriving on that page, a person might well become lost, and unable to make their way back to where they wanted to be.

SSI4 (male, 82) offered that seniors needed to read everything on the government webpages they accessed. He further advised seniors to:

> Make sure you understand what they are saying, not what your idea is, because it's easy to get confused, I'm talking about older people you know, they think they know what's going on, that's the only thing I can say, to educate the people to read what they're looking at.

It was his opinion that website designers didn't go out of their way to make them difficult to understand, but perhaps they assumed that seniors know what they're doing on the internet. He felt it wasn't always the case that seniors understood, and as a result, seniors make mistakes through not reading fully. His advice was directed at seniors to draw their attention to the need to read deeply on the sites.

It was SSI1's (female, 64) opinion that seniors were terrified of computers if they hadn't had much experience with them. While suggesting to another group of seniors that they participate in this study, she received feedback from them that: "I can't be bothered, it's too hard. I can turn it on and that's pretty well all I can do". SSI1 felt this was suggestive of a certain mindset, of an unwillingness to adapt to something new, a refusal to learn. As she noted, computers have been around for a few years now, and she felt that with the exception of seniors in their 80's, most weren't that old when computers became commonly used.

SSI5 (female, 68) also felt that seniors would have a difficult time with the government websites if they didn't have many computer skills or experience with them. She also expressed a security fear about unfamiliar websites, and thought that maybe other seniors were fearful of computer use for security reasons. SSI6 (female, 62) had observed that many seniors were just learning how to use computers, and were somewhat tentative in their use, but she didn't have any specific ideas about how seniors might use government websites. SSI7 (female, 61) had made similar observations about seniors attending computer instruction classrooms in the same room the study was undertaken. It was her opinion that age was a barrier to learning, as she put it, in regards to the computer training:

> They're short classes, only six weeks, free, an hour long, but some people that are in the beginners' class, they've been in the beginners' class for like five years. Some people they just can't, if it's not from their generation, they find it harder I think, and the instructor starts off very basic, just doing emails to begin with, because that's how seniors want to keep in touch with their children and grandchildren.

This comment shows a motivation for learning to use computers, something also noted by SSI8 (female, 69). She thought that computers were a great device for what she called 'shut-ins',

seniors who spent most of their time in their residence due to mobility and other issues. Like other participants, she felt that seniors were afraid of learning how to use computers, but insisted that they were a wonderful item for keeping the mind active, and conversing with people at distances. SSI9 (female, 65) somewhat summarized the comments with her opinion that websites needed to be user friendly. As she remarked,

> Some of them don't even use the internet so, to start, it just has to be user friendly, things that they can find quickly. The search engine, there's a search key in there, find out how to use that, once you know how to use that you can find it. So the search engine was actually a good help, trained to if you can't find it on the side menu bar then go to the search, probably what they need to be trained to be able to do

Summary of the Findings

This chapter has presented the quantitative and qualitative findings of the study as well as a description of the participants in the study. Participant descriptions were provided for both the main study (N=34) and the supplement interview (N=9). Quantitative findings were presented for four usability concepts – efficiency, effectiveness, satisfaction, and accessibility. The measures used for each of these concepts were stated, and numerical results were presented. Correlation charts were displayed to show relationships between the variables of completion rate, search rate, time, SUS scores, and age. The qualitative findings were presented within six thematic groups – adaptation, feelings, technical issues, navigation, evaluation, and suggestions. The themes were discussed using quotes from participants to describe the theme and support the argument for it.

Chapter 5 – Discussion

Introduction

In this chapter I discuss the findings from the study. I address the findings from the perspectives of the research questions and the framework and methods used. The chapter has two main sections. The first section discusses the findings and the answers that are provided to the individual research questions. This section also discusses the implications of the findings for the literature in the field. Section two reflects on how the theoretical framework and the data collection methods impacted the findings. I conclude this chapter by summarizing how my findings support and/or challenge the existing literature on the topic and evaluate the choice of theoretical framework and data collection methods in terms of what findings were made/not made in the study.

Discussion of the Research Questions and Findings

This study explored the user experience of seniors while performing information-seeking tasks on government websites tested for compliance with accessibility standards. This exploration constitutes the central questions in this study. In order to address it, a mixed-methods approach was employed. The approach included the collection of quantitative data through the numeric measures of the usability experiences of the participants and numeric measures of the accessibility of the websites involved. I also collected qualitative data in the form of narrative data through interviews and observations of the participants. SPSS and NVIVO software were used to assist in the analysis of the data.

There were two primary research questions. Two secondary questions were included within the first question:

1. What is the user experience of seniors when using government websites tested for compliance with accessibility standards?
    a. What barriers, if any, do seniors experience when using government websites?
    b. How do seniors adapt to the environment presented by government websites while using them?
2. What are the relationships among the user experience of seniors, barriers experienced, adaptations made (when using accessible government websites), and the reported accessibility rating (of the government websites)?

I will address each question separately showing how the findings inform our understanding of each one.

Research question #1.

What is the user experience of seniors when using government websites tested for compliance with accessibility standards?

This question starts with the concept of experience as previously discussed in the literature review. The literature suggested that experience had two primary criteria in the Dewey (1938/1997) schema, that of continuity and interaction. In this scheme, experience includes the effects of past experience on the present, and suggests that the present in turn influences the future, while functioning as the transactions between an individual and the environment. These concepts are displayed in Figure 3 below. I have included the term "time" to emphasize the progression of time through experience from past to future.

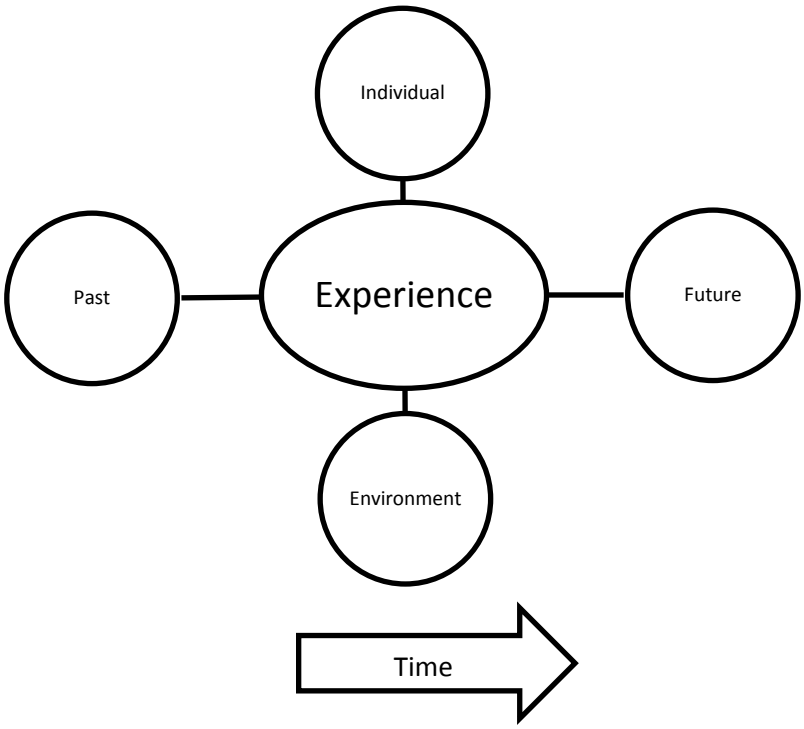

Figure 3. The components of experience, adapted from Dewey 1938/1997).

I employ these concepts within the digital technology environment or space used in this study. As such, experience within the bounds of the study might be thought of as the transactions between an individual and the digital website environment, including previous and future experiences with it. In a broader sense, user experience (UX) includes traditional usability measures (Rubin & Chisnell, 2008). It seeks to further expand the measures of the transactional experiences of the individual with the digital environment to include affective measures such as the hedonic quality, emotions, etc. (Hassenzahl & Tractinsky, 2006).

Both quantitative and qualitative data helped to address this research question. Accessibility measures were also explored and collected through automated accessibility checkers. Usability measures are addressed in this study through typical quantitative measures relating to efficiency, effectiveness, and satisfaction. Affective measures, and other usability

concepts such as learnability and usefulness (Rubin & Chisnell, 2008) were collected through the qualitative data. As such, all data collected within the study address the central question. Questions 1(a), 1(b), and question 2 address the specific experiences of the participants and serve to answer the central question, as discussed below. The accessibility data helped to describe the digital environment within which the study was situated. It addresses the central question directly by providing information as to the technical accessibility of the government websites used in the study.

The accessibility data were collected through two automated software tools, WebAim (2015) WAVE and AChecker (2015). The reported WAVE accessibility errors ranged widely within the four tested websites (see table 17 page 77, chapter 4). The Consumer website had the lowest mean and range of reported WAVE errors, while the Transport website had the highest mean and range of reported WAVE errors. For AChecker reported known problems, Consumer had the smallest range, Health had the lowest mean, and Labour had both the highest mean and highest range. These primary measures from both WAVE and AChecker were significantly and strongly correlated, suggesting that the automated checkers were measuring similar accessibility problems on the websites. However, it is also noted that different tools may have different coverage and interpretations of the guideline, producing different reported results on the same website (Vigo & Brajnik, 2011).

As noted earlier (p.82), the findings of the accessibility checking showed that all websites had known accessibility errors within as reported by the automated accessibility checkers and none were achieving full compliance with the web content accessibility guidelines (WCAG 2.0) level AA (W3C WCAG, 2008). The Ontario Regulation 191/11 mandates full conformance to these guidelines for Government of Ontario websites, phased in over a period of time. New

Government of Ontario websites (and their content) were required to comply by January 1, 2012, and all Government of Ontario websites and content must comply by January 1, 2016. This study shows that there is some work yet to be done to meet these requirements. The digital environment for the study can be characterized as mostly technically accessible while not achieving full guideline compliance. I return to the central question in the conclusions chapter six.

Research Question #1(a)

What barriers, if any, do seniors experience when using government websites? (Barriers within this environment - BWE).

The question of the barriers experienced by seniors when using government websites was addressed through the qualitative data. Barriers to technology usage by older persons were earlier described as being either self-imposed or societal-imposed due to misconceptions (myths) about their abilities (see Wandke et al., 2012). Participant observations during the task performance, along with utterances and answers given by the participants highlighted areas of concern. Three of the thematic categories shed light on the barriers experienced. These categories were discussed in the findings under the labels of feelings, technical issues, and navigation. This question of barriers is answered by these themes.

The participants' feelings about the use of the study websites were grouped into two forms in the discussion chapter. These forms were both negative in tone, and contributed to or were a result of their lack of success or pleasure in completing the tasks. The participants expressed frustrations with the websites in particular, and expressed questions of self-confidence about their internet skills in general. I differentiated participant comments between feelings and

evaluation, and noted that the emotional expressions were obvious only when they were negative in tone – none of the participants expressed positive emotions about the websites. Any positive utterances made were evaluative in nature – the website was good or bad, easy or not easy. In general, the emotional utterances contained the words 'I feel', or 'I don't feel', while going on to describe this feeling.

This above point seems quite important, and suggests that the websites examined are not of the character to effect positive emotions, just negative ones, with the exception of a sense of satisfaction occasionally expressed upon the completion of a task. The negative emotions observed were directed both at the websites and at the self. The websites either caused the participant to feel inadequate after attempts to use them, or the participant felt inadequate prior to using them and the website served to affirm their perceived inadequacy. In other words, a participant might blame the inability to complete tasks on the inadequacies on the website, or blame their own inadequacies for their inability to complete the tasks. Either of these two findings can be considered to be barriers to access. In the case of the former, the websites are not addressing the needs of the user. In the case of the latter, the users are not addressing the requirements of the websites.

The first case seems to be easily understandable. A solution to it would likely rely on design modifications, taking into account the needs of the user. The second case seems somewhat more problematic, as it suggests that the user must be made to accommodate the design of the websites. The literature recommends design modifications to address the particular needs of seniors (e.g. Money et al., 2011; Patsoule & Koutsabasis, 2014), while acknowledging website access barriers experienced by seniors.

The technical issues presented obvious barriers to website access. As noted in the findings, two primary issues arose during the study. The first issue I have called ‘slow computer’, and the second issue was that of colour contrast on webpages on the Ministry of Health website. The first issue involved the slow loading of pages on the Ministry of Transportation website, and, in the extreme, for six of the participants the pages would not load at all.

I learned that this was not an uncommon problem for the participants. Several reported that the same problem had occurred on their home computers. Their response in those cases was to abandon whatever tasks they were attempting to perform at the time, and to seek alternate ways to complete them, for example, by directly accessing a physical government services location. During the study, the participants responded to this problem in the same way, abandoning the task and noting that the barrier was a good reason to interact with government services at physical locations and not via the internet websites. This barrier served to reduce participants’ confidence in government websites.

Colour contrast on the Health website caused participants to have difficulty in locating the form for renewing an OHIP health card. Most of the links on the page were a light green in colour, and to some degree indistinguishable from one another. Participants remarked that this was not proper, in that they felt that it should be easy to distinguish and perceive what they were looking for. In addition, the page on which most had difficulty was noted to be very cluttered, with many different links and options, causing participants to delay in their choices and necessitating a thorough reading of the page that they felt was unnecessary and undesirable.

Navigation issues concerned reports of being lost while conducting their information searches. This issue arose from finding the desired information and not knowing where to navigate to next. Connected with this was the reported issue of too much information being provided on webpages/websites. Both of these issues can be considered as barriers to the usage of websites. In the first case, inability to locate information and the sense of being lost in navigation is a potential barrier to being able to successfully complete a desired information-seeking goal. Three participants that became lost were unable to complete the study tasks. I compared the cases of participants with very low task completion rates (less than half of the tasks completed successfully, N=6) and examined their comments. The comment regarding too much information was widespread, noted by twelve participants overall, with only two of this number also having a very low task completion rate. A sense of being lost appeared in the comments of eight participants, with three of this number having a very low task completion rate.

In a similar manner, too much information presented on a webpage contributed to the sense that finding the required information was either not possible or required more effort than participants were willing to expend. Participants wanted to be able to find information in a reasonably efficient and timely manner. If this was not possible, participants often would give up the search, with the exception of two participants (M33 and M1) who took the most time recorded. These participants were determined to complete the tasks, and were successful in completing two out of four, and three out of four, respectively.

A comparison of the time to complete the tasks with the qualitative data showed that the participants reporting too much information and/or a sense of lostness took varying amounts of time to complete the tasks. Complaints in this regard included the use of language that was not clear to the readers. The suggestion was made that websites for seniors could target that group in

a better fashion, using language that was more appropriate to them and removing information that would likely be of no use or benefit to them. In other words, target websites to be more topical, relevant, and appropriate to seniors, if the goal is to have them use them. Comments of an apologetic tone were made by participants, as to why this suggestion had not already been implemented. Participants allowed that the subject-matter of the websites might be such that complex language was required. It was also commented that perhaps the writers of the website content used complex language without thinking about the comprehension level or needs of the readers.

The navigation issues overlap with the technical and participant feelings in producing barriers to the usage of these websites. For example, when navigating to the point of being lost, participants began to report feeling inadequate, in some cases, blaming the lostness on their own inadequacies, and further reinforcing their lack of confidence in themselves. For those that claimed there was too much information available, the feeling was that the websites were to blame, especially in the case of language that was deemed to be unnecessarily complex and lengthy. The technical issues that I have reported on can be linked to, or seen as a consequence of, the issue of too much information, in that the participants' feelings were directed negatively towards the websites, and not towards themselves. In summary, the barriers to usage of these websites can be described as having two primary causes – problems perceived to be internal to the website design and problems perceived to be internal to the participant.

The problems perceived to be with the websites included: colour contrast, slow computer, and too much information. These problems help to create a sense of frustration with the websites. In the literature, these types of problems are argued as barriers to website access, and are

described and argued as including navigation issues, clutter, distractions, excessive functionality, and technical jargon (e.g. see Sayago and Blat, 2010; Sharit et al., 2011).

The problems perceived to be with the participants included: being lost and a lack of self-confidence. These problems help to create a sense of frustration with the self. In both cases, this sense of frustration renders the user experience negative (see Hassenzahl et al., 2013). It also contributes to a possible negative perspective on the future use of these websites (continuity of experience, see Dewey, 1938/1997). Is this in the best interests of the consumers (user) and the producers of government websites? The next research question addresses some of the adaptations made by participants to the difficulties and barriers experienced by them during the study.

Research Question #1(b).

How do seniors adapt to the environment presented by government websites while using them? (Adaptation within this environment - AWE)

Previous experiences with computer use helped participants adapt to a digital website environment. Participants reported having learned to use computers during their working careers, in some cases being early adopters of the technology and growing with the changes in technology over the course of many years. Other participants noted that experience in the subject matter of the information seeking tasks also helped in their ability to find the information. In addition to experience, participants reported that persistence was a key to their success. Persistence involves a ‘no-quit’ attitude, where the participant would keep at a task until they had achieved success (Merriam-Webster (2014b). Persistence included the adoption of learned adaptations. On the other hand, a group of participants reported giving up easily and finding

other methods to achieve the desired results. This in itself could be considered an adaptation to the environment, prompting the search for and selection of a different environment.

In the cases involving persistence, participants used exploring strategies that they had learned and adapted for themselves. The two main strategies were what I have called the reading strategy and the searching strategy. The reading strategy involves more time and effort for a participant. In this strategy, the participant reads and/or scans the content of the webpage, choosing to move forward through links on the page. As noted in Romano Bergstrom et al. (2013), website navigation poses special challenges for the older user, particularly when webpages are cluttered and contain distracting information that attracts attention but does not provide the information sought. On the other hand, the searching strategy is less time-consuming and more efficient. It involves entering keywords into the search box on a webpage and navigating from the results.

Byun and Finnie (2011) argued that, in general, government websites contain a large volume of information, making them difficult to navigate and leading to frustration on the part of older users. Chin and Fu (2012) argued that older users used more focused search strategies than younger adults. The younger adults were more exploratory in nature, while the older adults benefited by a reduction in the number of links required to access information. Chin and Fu (2012, p. 3039) found that younger adults "…were more exploratory, as they clicked more links, but visited and left a page more quickly, than older adults", whereas older adults "…did more focused search, as they clicked fewer (but more relevant) links and spent longer time deciding on a link". My findings are consistent with these studies, in that the search strategy I observed utilized fewer links and produced more efficient results that what I have termed the reading strategy.

For those participants who prefer to adapt by choosing another environment, one adaptation was to use a different interface. A commonly reported different interface was that of the tablet/smartphone. These machines have touch screen and voice recognition abilities that make their usage easier than searching on a desktop interface. Participants in particular reported satisfaction with the voice recognition ability, using it to locate websites easily.

Adaptation, as discussed in the findings, seems to be primarily a function of the will of the participant - the desire of the participant to persist within the digital environment. For those participants motivated to use the digital environment, persistence enables them to find ways to adapt themselves in order to achieve their goals. As discussed in Wagner, Hassanein, and Head (2010), motivation is linked to self-efficacy, and a higher level of self-efficacy results in higher computer use. Persistence then has a motivational component, suggesting a further relationship to the goals and the self-efficacy of the user. Self-efficacy in my study was revealed in comments which I have described under the theme of feelings, specifically as self-confidence.

One of the themes I found in the participant narratives was that of evaluation. Participants freely offered their evaluation of the websites. These evaluations included comments about the perceived usability of the websites, motivations and needs of the participants. In particular, I observed that participants had a need to be perceived as competent, as expressed by comments that extolled their abilities with computers and website. This apparent need is something that has been identified in the literature as a desire for competency, among other human needs. For example, earlier (page 33) I produced Hassenzahl's (2008) primary definition of user experience. He refines it further,

> Thus, the second part of my definition of UX states: *Good UX is the consequence of fulfilling the human needs for autonomy, competency, stimulations (self-oriented) relatedness, and popularity (others-oriented) through interacting with the product or service (i.e. hedonic quality). Pragmatic quality facilitates the potential fulfilment of be-goals* (Hassenzahl, 2008, p. 12).

Additionally, motivation to use the government websites was related to their perceived usability. Framed negatively, participants reported not being motivated to use websites perceived to be difficult to use. Another reported usability perception relates to a general dislike for computers. In this scenario, participants are not motivated to use computers/websites due to previous bad experiences with them.

The interview question requesting participant suggestions provided two related sub-themes. The first one concerned general suggestions for government websites, and the second concerned specific suggestions for seniors using government websites. In the first sub-theme, participants suggested that government websites should be made easier to use in general. The second sub-theme specifically made reference to the special needs and possible limitations of the senior population.

These two themes are also found in the extant literature. In addition to the concerns noted in Byun and Finnie (2011) about the difficulties navigating government websites that contain large amounts of information, Chou et al. (2013) argued that websites must be specifically designed for older users, taking into account their special requirements. These themes are also argued in the work of Hanson (2009, 2011), and Norval et al. (2014), among others.

Additionally, Sharit et al. (2011) have recommended avoiding the use of technical jargon in order to ensure webpages are more understandable for older users.

Research question #2.

What are the relationships between the user experience of seniors, barriers experienced, adaptations made (when using accessible government websites), and the reported accessibility rating (of the government websites)? (Relationships within this environment)

Qualitative and quantitative data were compared to address this question. Quantitative findings were compared to the qualitative findings to see if connections existed and whether the data sets corroborated each other. For example, I examined a quantitative finding of a correlation between accessibility contrast errors and task time taken against the qualitative data, and I found that participants referenced contrast issues in their comments.

WCAG 2.0 asserts that guideline compliance will enhance accessibility for persons with disabilities, and make "…web content more usable to users in general" (W3C WCAG 2008), by adherence to the four guiding principles that require content to be perceivable, operable, understandable, and robust. In addition, adherence to WCAG 2.0 assists older users as they may possess functional limitations due to age, including vision decline, hearing loss, motor skill diminishment, and cognitive effects (W3C, 2008). My study collected self-reported ratings on a seven-point Likert scale of vision, hearing, memory, movement/mobility, and the absence of pain. The study group mean scores for each of these items was above a value of five on a seven point scale, suggesting that, as a whole, functional limitations were not large issues for the participants in the study. This observation is supported by the absence of findings of significant correlations between the items and the measures of efficiency (time) and effectiveness

(completion rate) with two exceptions. Movement/mobility was significantly and positively correlated with the number of completions on the Labour website, r = .339, 95% BCaCI [.110, .546], p = 0.050. Pain-free status was negatively correlated with the time taken on the Health website, r = -.524, 95% BCaCI [-.821, -.017], p = 0.004. I made forty comparisons (five health items against 4 time measures and 4 completion measures) in this specific analysis, and only two significant correlations were found. As a result, I don't place much emphasis on the significant findings and conclude that functional limitations were not large issues in this study, particularly in the absence of any participant utterances or observations about limiting factors of pain or movement during their task performances. However, while this particular study group was able to access the websites, and did not suffer from obvious functional limitations, the data suggest that other groups of seniors may have more limited functionalities (HRSDC, 2009). In light of these latter facts, it is certainly desirable that full compliance with WCAG 2.0 is achieved.

One accessibility issue that is of concern in the study is that of contrast errors reported by WAVE. Participants had noted colour issues on the Health website, as reported in the qualitative findings. Their comments suggested that the lack of contrast amongst items on the relevant webpages on that site deterred them from locating the information they sought. An analysis showed that the time spent on the all websites (N=28, Transport problem sites removed for this analysis) was significantly and positively correlated to the number of contrast errors reported on the sites and webpages, r = .435, 95% BCaCI [.052, .698], p = 0.021. This finding suggests a strong correlation between contrast errors and time taken, corroborating the statements made by participants during their task performance. Although the relationship can't be determined to be causal in this study, it does speak to a significant number of contrast issues prevalent. It might be an expected result given that time taken also correlates positively with the number of pages

accessed. However, considering the assertion that vision declines with age (W3C, 2008), this finding emphasizes the need to provide proper contrasting on webpages and content in order that older users can efficiently perceive and thereby process the content. It is further emphasized by the fact that the participants self-reported a lack of problems with their vision, meaning that a group without vision problems struggled with the contrast issues.

The time spent on performing the information-seeking tasks on the government websites was used as a measure of efficiency. For three websites, the mean time taken was higher for incomplete tasks than for completed tasks with the mean time taken being approximately the same for incomplete and complete task performance on the Labour site. The range, minimum, maximum, and means (in seconds) are shown in table 23 for all time taken on the individual websites, whether the task was completed or not:

Table 23.

Time taken (in seconds) on tasks whether completed or not.

	N	Range	Minimum	Maximum	Mean
Health	34	525	45	570	176.18
Transport	28	575	40	615	169.29
Labour	34	700	50	750	225.44
Consumer	34	300	30	330	154.71
Valid N (listwise)	34				

This table shows that participants spent the lowest mean time on the Consumer website and spent the highest mean time on the Labour website, with a wide range of times taken, particularly on the three websites (Health, Transport, and Labour) that had the highest mean times.

The effectiveness component of usability was measured as the completion rates for the tasks on each website. In my study, the completion rate was highest for the Consumer site, and lowest for the Labour site, as measured individually by website. The learnability component of usability is a part of the effectiveness component and is concerned with the ability of a user to become competent with a product over time (Rubin & Chisnell, 2008). In the qualitative data of my study, this learnability component can be observed within the theme of adaptation reported in the findings. The findings within this theme note that adaptation and learnability are gained through previous experience, persistence, adopting strategies, and using different interfaces to gather the required information.

Satisfaction was measured primarily through the System Usability Scale (SUS). The results of this measurement were that participants were overall least satisfied with the Labour site and most satisfied with the Consumer site. These measures were corroborated in the qualitative data, as reported by participants and gathered under the theme of evaluation. In this theme, participants reported widely that websites were either easy or not easy to use, along with reporting on perceived usability and their motivation to use the websites. Both perceived usability and motivation are components of the usefulness of the websites (Rubin & Chisnell, 2008). In my study, participants were divided on these elements, some not finding the government websites useful enough to them to spend the time learning how to use them. Some of these participants were dissuaded by the use of computers in general, lacking motivation to learn how to use them better, if at all. Other participants were motivated to use computers by a desire to become competent in their use in order to facilitate social interaction. And these latter participants also reported a desire to be perceived to be competent, to not be left behind in the digital and interconnected internet world.

These observations around motivation and perceived usability connect with the descriptions of do-goals and be-goals given by Diefenback & Hassenzahl (2010). In their description, do-goals possess pragmatic quality, and be-goals possess hedonic quality (see also Hassenzahl, 2008). Pragmatic quality refers to the perceived ability of a product to satisfy the achievement of a task, while hedonic quality refers to the perceived ability of a product to satisfy self-needs such as competence, self-expression etc. (see also Brajnik & Giachin, 2014). Diefenback & Hassenzahl (2010) make the argument that user needs must be examined beyond the pragmatic, and warn that all too often the justification for the choice or design of a product is based on pragmatic quality, when users are at least as interested in the hedonic quality. As Rubin & Chisnell (2008) also note, usefulness is about the willingness to use a product, and:

> If a system is easy to use, easy to learn, and even satisfying to use, but does not achieve the specific goals of a specific user, it will not be used even if it is given away for free. Interestingly enough, usefulness is probably the element that is most often overlooked during experiments and studies in the lab. (pg. 4)

Similarly, Brajnik and Giachin (2014) argue that user experience combines hedonic and pragmatic qualities and goes beyond mere usability and accessibility. They argue that user experience should include the idea that experiences are best measured not only in terms of their usefulness in achieving tasks, but also in their usefulness in satisfying social, psychological, and physiological needs. In my study, the participants were all able to access the websites, but to what extent could they use them, and what was the quality of their experience? The secondary questions helped in addressing these issues, through the grouping of the participant responses and observations into the themes of adaptation, feelings, technical issues, navigation, evaluation, and suggestions. I've argued that the quality of the participants' experience included a sense of

frustration with the websites, along with perceptions and evaluations about the usability of them. I've argued that the quality of their experience is holistic in nature, encompassing the qualities that Brajnik and Giachin (2014) note.

In summarizing this section, both the perceptions of the usefulness of the websites, and the experience quality varied widely. Common threads in this variance include the rejection of the websites by some participants, the suggestions to improve the quality by giving more consideration to the needs of seniors, contrasted with the report that the websites weren't bad once one got used to them and/or developed some necessary skill level to navigate them.

Impact of the Theoretical Framework and Data Collection Methods on Findings

The theoretical framework included conceptual elements from the literature on seniors' characteristics, usability research, and user experience theory. In using these elements, the data collection methods were designed to record measurables (QUAN) of a number of the concepts used. In particular, drawing from the usability research literature, I employed the concepts of accessibility, satisfaction, effectiveness, and efficiency. I expanded from these usability concepts to an argument in favour of adopting user experience concepts found in the research literature. The user experience concepts included hedonic and pragmatic quality, perceived usability, and be-goals and do-goals. I further argued for a connection between the theory of experience authored by Dewey (1938/1997) and the user experience theory of Hassenzahl (2014), drawing parallels between the concepts of interaction, continuity, and meaningfulness as expressed by both authors in their respective eras.

These theoretical concepts helped to tailor the data collection methods used. In employing theoretical usability concepts, I also employed standard usability collection methods

for acquiring quantitative data (time, completion scores, heuristic accessibility scores, and a satisfaction scale, for example see Sauro & Lewis, 2012). In order to attempt to understand user experience concepts for this study, I employed qualitative data collection methods (interviews and participant observations, for example see Brajnik & Giachin, 2014). These differing and multiple methods of data collection provoked the development of a mixed-methods research design.

The mixed-methods research design impacted upon the findings by delivering a wealth of data in multiple forms. The quantitative and qualitative data produced served to provide a means for cross-referencing the findings to assist with determining a measure of validity and reliability, along with permitting an integration of the two streams of data collected. In this manner, the study provides more than one perspective on the substantive topic. Both objective and subjective findings are incorporated into the overall findings that help to provide a complex and compelling argument supporting the conclusions made in the next chapter.

Summary of Discussion of Findings

The findings in my study support previous literature and what I think of as a movement towards a more qualitative and subjective turn towards determining usability and user experience. In this qualitative turn, the idea of hedonic quality as contrasted yet co-existent with pragmatic quality are emphasized (Diefenback & Hassenzahl, 2011). In some cases in the literature, attempts are made to measure the hedonic constructs explicitly (e.g. Brajnik & Giachin, 2014). In other cases, the qualitative emphasis takes the form of ethnographic methods for studying seniors using the internet (e.g. Sayago & Blat, 2010). Regardless of the form employed, recent literature in the field suggests multidisciplinary methods are beneficial (Silva,

Braga, & Teixeira, 2014) and that a shift from quantitative to qualitative and mixed methodologies has occurred (Bargas-Avila and Hornbaek, 2011).

In the next chapter I will report on my conclusions about how the findings answer the central question of the user experience of seniors in this study. In doing so, I will draw on the discussion of the answers to the secondary research questions, using that to make conclusions about the central question.

Chapter 6 – Conclusions and Implications

This chapter discusses the conclusions I have reached from the research findings. I relate these conclusions to the literature, referring back to the literature review. I reflect upon and evaluate how the findings address the central question and aims of the study. I conclude with recommendations on how to utilize this study's findings and for future research.

In the literature review I discussed the use of narrative as a way to understand the experiences of participants in my study. I suggested that the narratives of the participants were both explanatory of their experiences, and representative of participants' thoughts, feelings, and attitudes. The value of viewing participant statements as narrative was emphasized to me as I conducted the analysis of the participant statements and utterances. These narratives served to situate the participants within a context. This context incorporated their experiences with the use of technology and websites.

The research findings in my study suggest that more attention needs to be directed at satisfying the needs of seniors using government websites if an increased uptake in the use of these websites is desired. The claim that attention should be directed more at satisfying user needs is further discussed in the literature by authors such as Cooper et al. 2012, Newell 2008, Law 2010, among others. I have suggested that these needs are at least two-fold in nature, borrowing from the thinking of Hassenzahl (2008) on the nature of be-goals and do-goals. In the Hassenzahl scheme, be-goals should be emphasized as they meet the subjective needs of a technology user. Be-goals involve the human feelings such as competence and stimulation, while do-goals involve the human ability to achieve tasks. In the support of task achievement, do-goals help but are not sufficient to fulfill be-goals. It is my claim that the findings in my study support

the idea of emphasizing the be-goals of seniors using government websites specifically, and websites and the internet in general terms.

Earlier I discussed the need to dispel the myths and stereotypes surrounding the older technology-user. Fairweather (2008) recommended that assumptions about user groups should be avoided. Wandke et al. (2012) set out to directly dispel six myths about older people using computers, while acknowledging older people have differing attributes from younger people. Arguments of this nature suggest that age-related barriers might better be thought of as (senior) population characteristics. If barriers are thought of as myths, and characteristics as reality, then perhaps it is easier to take a different approach to the design situation from the start, as suggested by Silva et al. (2014). Casting the issue in this language removes the negativity suggestive in the idea that barriers are to be overcome. Instead, the idea becomes one of characteristics to be addressed.

Reframing the language used in studying seniors and their internet usage retains the onus on the product (websites) to be amenable to the user (seniors). Perhaps it even positively reinforces the idea behind this onus – that government products are meant to serve the public. In this study it was the case that websites have achieved or were close to achieving an acceptable level of accessibility compliance. The findings of unsatisfactory usability and experience leave the conclusion that the problem still lies with the product, even after the accessibility initiatives. For example, study participants enumerated task-related goals for internet usage (e.g. banking, shopping, research, etc.) but also made mention of other uses and goals. These other goals included the use of the internet for pleasure, social networking, and entertainment. These uses might well be characterized as less task-related (do-goals) and more stimulative (be-goals) in nature. Combined with these uses were such things as the apparent desire by participants to

appear competent to others, along with a need to feel confident in ones' self. I think of these latter goals as *experiential quality*, in keeping with both Dewey's (1938/1997) and Hassenzahl's (2014) ideas on the quality of experience (see Table 2, page 34).

The quantitative findings help to describe the user experience of the participants in showing relationships between completion rate and age, for example (see table 22, page 80). The finding shows an inverse medium correlation (r = -.388). The inverse relationship showed that as participant age increased, the task completion rate decreased. While at first glance this finding might support a conclusion that some of the myths about older people are in fact true (see Wandke et al., 2012), a closer look at the qualitative data helps to contextualize this finding. For example, three participants (ages 74, 82, and 84) failed to complete any of the four tasks satisfactorily. In each case, these participants expressed a lack of interest in using the government websites, making comments about the websites containing too much irrelevant information and being difficult to navigate. These comments suggest a lack of motivation to complete the tasks and use government websites. In each case, these participants reported being daily internet users of other services and information. This type of contradictory finding suggests an avenue for further research. In my study, I did not pursue this as thoroughly as might be done, and suggest that the topic would benefit from further research in this area. Further research could attempt to better isolate the motivational factors and compare them more thoroughly to completion rates. Some preliminary work has also been done in this regard (e.g. Sayago, Sloan & Blat, 2011).

Future research could also be directed at investigating the characteristics of seniors' internet needs and wants, specifically regarding government websites but also towards seniors' internet use in general, within a paradigm of *experiential quality* in the framework of user

experience. An expansion of the sample size would yield more generalizable results. The topic would also benefit through the study of other groups of seniors, including variables such as cultural diversity and social economic status.

Coincidentally, as this dissertation paper was being drafted, the Apple Corporation (2015) issued a press release detailing a cooperative effort with IBM and a Japanese government-owned corporation (Japan Post Group) to issue iPads and special applications to seniors in Japan. The purpose of this initiative is to improve the quality of life for Japanese seniors by providing technology to permit better connections (experiences) between seniors and the broader society. The press release describes the aging of the Japanese society, and the near future projection of a population containing 40% seniors. The plan is to provide iPads and applications to over 4 million Japanese seniors in the next five years. The provided technology will assist seniors with their specific needs, through the use of custom-built apps and accessibility features. As Tim Cook, Apple's CEO remarks in the press release:

> This initiative has potential for global impact, as many countries face the challenge of supporting an aging population...iPad is incredibly intuitive, easy to use and has accessibility features built in, making it a perfect device for any generation to be connected and engaged (Apple Corporation, 2015, pg. 1).

Recommendation

With this latest initiative in mind, I recommend that (Canadian) government authorities seek to improve the experiential quality for seniors interacting with the internet. The recommendation also suggests the desirability of improved experiential quality as a goal of

website design. My research suggests that government websites should support seniors by addressing the improvement of the user experience.

References

Accessibility for Ontarians with Disabilities Act, 2005. Statutes of Ontario, 2005, Chapter 11. Retrieved from http://www.e-laws.gov.on.ca/

AChecker. (2015a). *Web Accessibility Checker.* Retrieved from http://achecker.ca/checker/index.php

AChecker. (2015b). *AChecker Handbook.* Retrieved from http://achecker.ca/documentation/index.php?p=checker/index.php

Affonso de Lara, S. M., Watanabe, W. M., dos Santos, E. P. B., & Fortes, R. P. M. (2010). Improving WCAG for elderly web accessibility. *Proceedings of the 28th ACM International Conference on Design of Communication,* São Carlos, São Paulo, Brazil. 175-182. doi: 10.1145/1878450.1878480

Apple Corporation. Apple Press Info (April 30, 2015). *Japan Post Group, IBM and Apple Deliver iPads and Custom Apps to Connect Elderly in Japan to Services, Family and Community* . Retrieved from http://www.apple.com/pr/library

Bargas-Avila, J. A., & Hornbaek, K. (2011). Old wine in new bottles or novel challenges: A critical analysis of empirical studies of user experience. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, Vancouver, BC, Canada. 2689-2698. doi: 10.1145/1978942.1979336

Bazeley, P. (2007). *Qualitative data analysis with NVivo* (2nd ed.). Thousand Oaks, CA: Sage.

Brajnik, G., & Giachin, C. (2014). Using sketches and storyboards to assess impact of age difference in user experience. *International Journal of Human-Computer Studies*, *72*(6), 552-566.

Brooke, J. (1996). SUS: A "quick and dirty" usability scale. In P. Jordan, B. Thomas & B. Weerdmeester (Eds.), *Usability evaluation in industry* (pp. 189-194). London, UK: Taylor & Francis.

Bruner, J. (1987). Life as narrative. *Social Research, 71*(3). Retrieved from http://www.jstor.org/stable/40970444

Bruner, J. (Autumn 1991). The narrative construction of reality. *Critical Inquiry,* 18.

Bruner, J. (1996). *The Culture of Education.* Cambridge, Massachusetts: Harvard University Press.

Burmeister, O.K. (2010). Websites for seniors: Cognitive accessibility. *International Journal of Emerging Technologies and Society,8*(2), 99-113. Retrieved from http://www.swinburne.edu.au/hosting/ijets/ijets/

Byun, D., & Finnie, G. (2011). Evaluating usability, user satisfaction and intention to revisit for successful e-government websites. *Electronic Government, an International Journal,* (8), 1-19. doi:10.1504/EG.2011.037694

Chen, J. V., Lin, C., Yen, D. C., & Linn, K. (2011). The interaction effects of familiarity, breadth and media usage on web browsing experience. *Computers in Human Behavior, 27*(6), 2141-2152. doi: 10.1016/j.chb.2011.06.008

Chin, J., & Fu, W. (2012). Age differences in exploratory learning from a health information website. *Proceedings of the 2012 ACM Annual Conference on Human Factors in Computing Systems,* Austin, Texas, USA. 3031-3040. doi: 10.1145/2208636.2208715

Chou, W. H., Lai, Y. T., & Liu, K. H. (2013). User requirements of social media for the elderly: a case study in Taiwan. *Behaviour & Information Technology*, *32*(9), 920-937.

Clandinin, D. J., & Connelly, F. M. (2000*). Narrative inquiry: Experience and story in qualitative research*. San Francisco, CA: Wiley.

Cohen, J. (1992). A power primer. *Psychological Bulletin, 112*(1), 155-159. doi:http://dx.doi.org/10.1037/0033-2909.112.1.155

Collins, K.M.T., Onwuegbuzie, A.J., & Jiao, Q.G. (2007). A mixed methods investigation of mixed methods sampling designs in social and health science research. *Journal of Mixed Methods Research 2007 1*(3). 267-294. doi:10.1177/1558689807299526

Connelly, F. M., & Clandinin, D. J. (1990). Stories of experience and narrative inquiry. *Educational researcher*, *19*(5), 2-14.

Cooper, M., Sloan, D., Kelly, B., & Lewthwaite, S. (2012). A challenge to web accessibility metrics and guidelines: Putting people and processes first. *Proceedings of the International Cross-Disciplinary Conference on Web Accessibility,* Lyon, France. 20:1-20:4. doi: 10.1145/2207016.2207028

Creswell, J.W. (2009). *Research design: Qualitative, quantitative, and mixed methods approaches* (3rd ed.). Thousand Oaks, CA: Sage.

Creswell, J.W., & Plano Clark, V.L. (2011). *Designing and conducting mixed methods research* (2nd ed.). Thousand Oaks, CA: Sage.

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.

Czaja, S. J., Gregor, P., & Hanson, V. L. (2009). Introduction to the special issue on aging and information technology. *ACM Trans.Access.Comput., 2*(1), 1:1-1:4. doi: 10.1145/1525840.1525841

Czaja, S., Sharit, J., Hernandez, M., Nair, S., & Lowenstein, D. (2010). Variability among older adults in internet health information-seeking performance. *Gerontechnology, 9*(1), 46-55. doi: 10.4017/gt.2010.09.01.004.00

Davies, M. B. (2007). *Doing a successful research project using qualitative or quantitative methods.* New York: Palgrave MacMillan.

Denzin, N.K., & Lincoln, Y.S. (2008). Introduction: The discipline and practice of qualitative research. In N. K. Denzin & Y. S. Lincoln (Eds.), *Strategies of qualitative inquiry* (3rd ed.). Thousand Oaks: Sage.

Dewey, J. (1938/1997). Experience and education. In A.L. Hall-Quest (Ed.), *Experience and education by John Dewey: The Kappa Delta Pi lecture series.* New York: Simon and Schuster. (Original work published 1938).

Dickinson, A., Arnott, J., & Prior, S. (2007). Methods for human-computer interaction research with older people. *Behaviour & Information Technology, 26*(4), 343-352. doi: 10.1080/01449290601176948

Dickinson, A., Eisma, R., & Gregor, P. (2011). The barriers that older novices encounter to computer use. *Universal Access in the Information Society, 10*(3), 261-266.

Diefenbach, S., & Hassenzahl, M. (2011). The dilemma of the hedonic–Appreciated, but hard to justify. *Interacting with Computers, 23*(5), 461-472.

Driscoll, D. L., Appiah-Yeboah, A., Salib, P., & Rupert, D. J. (2007). Merging qualitative and quantitative data in mixed methods research: How to and why not. *Ecological and Environmental Anthropology (University of Georgia),* 18. Retrieved from http://digitalcommons.unl.edu/icwdmeea/18

Fairweather, P. G. (2008). How older and younger adults differ in their approach to problem solving on a complex website. *Proceedings of the 10th International ACM SIGACCESS Conference on Computers and Accessibility,* Halifax, Nova Scotia, Canada. 67-72. doi: 10.1145/1414471.1414485

Field, A. (2013). *Discovering statistics using IBM SPSS statistics* (3rd ed.). Thousand Oaks, CA: Sage.

Fisk, A.D., Rogers, W.A., Charness, N. Czaja, S.J., & Sharit, J. (2009). *Designing for older adults: Principles and creative human factors approaches* (2nd ed.). Boca Raton, FL: CRC Press.

Frøkjær, E., Hertzum, M., & Hornbæk, K. (2000, April). Measuring usability: Are effectiveness, efficiency, and satisfaction really correlated?. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems* (CHI '00). ACM, New York, NY, USA, 345-352.

Gay, G., & Li, C. Q. (2010, April). AChecker: open, interactive, customizable, web accessibility checking. In *Proceedings of the 2010 International Cross Disciplinary Conference on Web Accessibility (W4A)* (W4A '10). ACM, New York, NY, USA, Article 23, 2 pages.

Hanson, V. L. (2009). Age and web access: The next generation. *Proceedings of the 2009 International Cross-Disciplinary Conference on Web Accessibililty (W4A)*, Madrid, Spain. 7-15. doi: 10.1145/1535654.1535658

Hanson, V. L. (2011). Technology skill and age: What will be the same 20 years from now? *Univers.Access Inf.Soc., 10*(4), 443-452. doi: 10.1007/s10209-011-0224-1

Hart, T. A., Chaparro, B. S., & Halcomb, C. G. (2008). Evaluating websites for older adults: adherence to 'senior-friendly'guidelines and end-user performance. *Behaviour & Information Technology*, *27*(3), 191-199.

Hassenzahl, M. (2014): User Experience and Experience Design. In Soegaard, M., and Dam, R.F., (Eds.). *The Encyclopedia of Human-Computer Interaction* (2nd ed.). Aarhus, Denmark: The Interaction Design Foundation. Retrieved from https://www.interaction-design.org/encyclopedia/user_experience_and_experience_design.html

Hassenzahl, M. (2008). User experience (UX): Towards an experiential perspective on product quality. *Proceedings of the 20th International Conference of the Association Francophone d'Interaction Homme-Machine*, Metz, France. 11-15. doi: 10.1145/1512714.1512717

Hassenzahl, M., Eckoldt, K., Diefenbach, S., Laschke, M., Lenz, E., & Kim, J. (2013). Designing moments of meaning and pleasure. Experience design and happiness. *International Journal of Design*, *7*(3), 21-31.

Hassenzahl, M., & Tractinsky, N. (2006). User experience - a research agenda. *Behaviour & Information Technology, 25*(2), 91-97. doi: 10.1080/01449290500330331

Hertzum, M., & Clemmensen, T. (2012). How do usability professionals construe usability? *International Journal of Human - Computer Studies, 70*(1), 26-42. doi: 10.1016/j.ijhcs.2011.08.001

Hollender, N., Hofmann, C., Deneke, M., & Schmitz, B. (2010). Integrating cognitive load theory and concepts of human-computer interaction. *Computers in Human Behavior, 26*(6), 1278. doi: 10.1016/j.chb.2010.05.031"

Hornbaek, K., & Law, E. L. (2007). Meta-analysis of correlations among usability measures. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems,* San Jose, California, USA. 617-626. doi: 10.1145/1240624.1240722

HRSDC. (2009). *Advancing the Inclusion of People with Disabilities 2009.* Human Resources and Skills Development Canada. Retrieved from http://www.hrsdc.gc.ca/eng/disability_issues/reports/fdr/2009/page05.shtml#_5.2

Jaspers, M. W. M. (2009). A comparison of usability methods for testing interactive health technologies: Methodological aspects and empirical evidence [Abstract]. *International Journal of Medical Informatics, 78*(5) 340-353.

Johnson, R. B., Onwuegbuzie, A. J., & Turner, L. A. (2007). Toward a definition of mixed methods research. *Journal of Mixed Methods Research, 1*(2), 112-133. doi: 10.1177/1558689806298224

Kim, J., & Han, S. H. (2008). A methodology for developing a usability index of consumer electronic products. *International journal of industrial ergonomics, 38*(3), 333-345.

Kincl, T., & Strach, P. (2012). Measuring website quality: Asymmetric effect of user satisfaction. *Behaviour & Information Technology, 31*(7), 647-657. doi: 10.1080/0144929X.2010.526150

Kirschner, P. A. (2002). Cognitive load theory: Implications of cognitive load theory on the design of learning. *Learning and Instruction, 12*(1), 1. doi: 10.1016/S0959-4752(01)00014-7"

Kurniawan, S., & Zaphiris, P. (2005). Research-derived web design guidelines for older people. *Proceedings of the 7th International ACM SIGACCESS Conference on Computers and Accessibility,* Baltimore, MD, USA. 129-135. doi: 10.1145/1090785.1090810

Kvale, S. (2007). *Doing interviews.* Thousand Oaks, CA: Sage.

Law, E. L. (2010). Modelling user experience – an agenda for research and practice. *Interacting with Computers, 22*(5), 313-322. doi: 10.1016/j.intcom.2010.04.006

Law, E. L. (2011). The measurability and predictability of user experience. *Proceedings of the 3rd ACM SIGCHI Symposium on Engineering Interactive Computing Systems,* Pisa, Italy. 1-10. doi: 10.1145/1996461.1996485

Lynch, K. R., Schwerha, D. J., & Johanson, G. A. (2013). Development of a weighted heuristic for website evaluation for older adults. *International Journal of Human-Computer Interaction, 29*(6), 404-418.

Merriam-Webster. (2014a). *Attitude.* Retrieved from http://www.merriam-webster.com/dictionary/attitude

Merriam-Webster. (2014b). *Persistence.* Retrieved from http://www.merriam-webster.com/dictionary/persistence

Money, A., Lines, L., Fernando, S., & Elliman, A. (2011). E-government online forms: Design guidelines for older adults in Europe. *Universal Access in the Information Society, 10*(1), 1-16. doi: 10.1007/s10209-010-0191-y

Meneweger, T., Wurhofer, D., Obrist, M., Beck, E., & Tscheligi, M. (2014, April). Characteristics of narrative textual data linked to user experiences. In *CHI'14 Extended Abstracts on Human Factors in Computing Systems* (pp. 2605-2610). ACM.

Money, A., Lines, L., Fernando, S., & Elliman, A. (2011). E-government online forms: Design guidelines for older adults in europe. *Universal Access in the Information Society, 10*(1), 1-16. doi: 10.1007/s10209-010-0191-y

Neilsen, J. (2012a). *Usability 101: Introduction to usability.* Retrieved from http://www.nngroup.com/articles/usability-101-introduction-to-usability/

Neilsen, J. (2012b). *How many test users in a usability study?* Retrieved from http://www.nngroup.com/articles/how-many-test-users/

Newell, A. F. (2008). Accessible computing -- past trends and future suggestions: Commentary on "Computers and people with disabilities". *ACM Trans.Access.Comput., 1*(2), 9:1-9:7. doi: 10.1145/1408760.1408763

Newell, A. F. (2011). Design and the digital divide: Insights from 40 years in computer support for older and disabled people. *Synthesis Lectures on Assistive, Rehabilitative, and Health-Preserving Technologies, 1*(1), 1-195. doi: 10.2200/S00369ED1V01Y201106ARH001

Norval, C., Arnott, J. L., & Hanson, V. L. (2014, April). What's on your mind?: investigating recommendations for inclusive social networking and older adults. In *Proceedings of the 32nd annual ACM conference on Human factors in computing systems* (pp. 3923-3932). ACM.

Obrist, M., Roto, V., Vermeeren, A., Väänänen-Vainio-Mattila, K., Law, E. L. C., & Kuutti, K. (2012, May). In search of theoretical foundations for UX research and practice. In *CHI'12 Extended Abstracts on Human Factors in Computing Systems* (pp. 1979-1984). ACM.

Ontario Ministry of Economic Development, Employment & Infrastructure. (2015a). *Understanding accessibility: Why accessibility is good for Ontario.* Retrieved from http://www.mcss.gov.on.ca/en/mcss/programs/accessibility/understanding_accessibility/index.aspx

Ontario Ministry of Economic Development, Employment & Infrastructure. (2015b). *A guide to the integrated accessibility standards regulation – April 2014.* Retrieved from: *http://www.mcss.gov.on.ca/en/mcss/programs/accessibility/other_standards/iasr_2012/toc_iasr.aspx*

Ontario Public Service (2011). *Ontario public service (OPS) accessible customer service policy.* Retrieved from http://www.health.gov.on.ca/en/

Ontario Regulation 191/11, (2011). *Integrated accessibility standards, made under the Accessibility for Ontarians with Disabilities Act, 2005.* Toronto, Ontario: Ontario Provincial Government. Retrieved from http://www.e-laws.gov.on.ca/html/source/regs/english/2011/elaws_src_regs_r11191_e.htm

Palys, T. (2003). *Research decisions: Quantitative and qualitative perspectives* (3rd ed.). Thousand Oaks, CA: Sage.

Patsoule, E., & Koutsabasis, P. (2012). Redesigning web sites for older adults. *Proceedings of the 5th International Conference on PErvasive Technologies Related to Assistive Environments,* Heraklion, Crete, Greece. 13:1-13:8. doi: 10.1145/2413097.2413114

Patton, M. Q. (2002). *Qualitative research and evaluation methods*. (3rd ed.) Thousand Oaks, CA: Sage Publications.

Richards, L. (2005). *Handling qualitative data: A practical guide.* Thousand Oaks, CA: Sage.

Righi, V., Sayago, S., & Blat, J. (2011). Towards Understanding E-Government with Older People and Designing an Inclusive Platform with them-Preliminary Results of a Rapid Ethnographical Study. *International Journal of Public Information Systems*, *7*(3).

Romano Bergstrom, J. C., Olmsted-Hawala, E. L., & Jans, M. E. (2013). Age-related differences in eye tracking and usability performance: website usability for older adults. *International Journal of Human-Computer Interaction*, *29*(8), 541-548.

Romen, D., & Svanaes, D. (2012) Validating WCAG versions 1.0 and 2.0 through usability testing with disabled users. *Universal Access in the Information Society,11*(4) 375-385. doi:10.1007/s10209-011-0259-3

Rubin, J., & Chisnell, D. (2008). *Handbook of Usability Testing: How to plan, design, and conduct Effective Tests (2nd ed.).* Indianapolis, IN: John Wiley & Sons.

Rugg, G., & Petre, N. (2007). *A gentle guide to research methods.* Berkshire, England: McGraw-Hill.

Sauro, J., & Lewis, J. R. (2011). When designing usability questionnaires, does it hurt to be positive? *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems,* Vancouver, BC, Canada. 2215-2224. doi: 10.1145/1978942.1979266

Sauro, J., & Lewis, J. R. (2012). *Quantifying the user experience: Practical statistics for user research*. Waltham, MA: Morgan Kaufman.

Sayago, S., & Blat, J. (2010). Telling the story of older people e-mailing: An ethnographical study. *International Journal of Human-Computer Studies, 68*(1-2). 105-120. doi: 10.1016/j.ijhcs.2009.10.004"

Sayago, S., Guijarro, J., & Blat, J. (2012). Selective attention in web forms: An exploratory case study with older people. *Behaviour & Information Technology, 31*(2), 171-184. doi: 10.1080/01449291003767920

Sayago, S., Sloan, D., & Blat, J. (2011). Everyday use of computer-mediated communication tools and its evolution over time: An ethnographical study with older people. *Interacting with Computers, 23*(5), 543. doi: 10.1016/j.intcom.2011.06.001"

Scherer, K. (2005). What are emotions? And how can they be measured? *Social Science Information, 44*(4), 695-729. doi: 10.1177/0539018405058216

Sharit, J., Hernandez, M. A., Nair, S. N., Kuhn, T., & Czaja, S. J. (2011). Health problem solving by older persons using a complex government web site: Analysis and implications for web design. *ACM Trans.Access.Comput., 3*(3), 11:1-11:35. doi:10.1145/1952383.1952386

Silva, S., Braga, D., & Teixeira, A. (2014). AgeCI: HCI and Age Diversity. In C. Stephanidis & M. Antona (Eds.), *Universal Access in Human-Computer Interaction. Aging and Assistive Environments* (pp. 179-190). Switzerland: Springer International Publishing.

Smith, P. A. (1996). Towards a practical measure of hypertext usability. *Interacting with Computers, 8*(4), 365-381. doi: 10.1016/S0953-5438(97)83779-4

Sørum, H., Andersen, K. N., & Vatrapu, R. (2012). Public websites and human–computer interaction: An empirical study of measurement of website quality and user satisfaction. *Behaviour & Information Technology, 31*(7), 697-706. doi: 10.1080/0144929X.2011.577191

Stewart, T. (2012). Websites – quality and usability. *Behaviour & Information Technology, 31*(7), 645-646. doi: 10.1080/0144929X.2012.707372

Subasi, O., Leitner, M., Hoeller, N., Geven, A., & Tscheligi, M. (2011). Designing accessible experiences for older users: User requirement analysis for a railway ticketing portal. *Univers.Access Inf.Soc., 10*(4), 391-402. doi:10.1007/s10209-011-0223-2

Sweller, J., van Merrienboer, J., & Paas, F. (1998). Cognitive architecture and instructional design. *Educational Psychology Review, 10*(3), 251-296.

Teddlie, C., & Tashakkori, A. (2009). *Foundations of mixed methods research: Integrating quantitative and qualitative approaches in the social and behavioral sciences*. Thousand Oaks, CA: Sage.

Urdan, T. C. (2010). *Statistics in plain English* (3rd ed.).New York: Routledge.

Veenhof, B., & Timusk, P. (2009). *Online activities of Canadian boomers and seniors.* (No. 11-008). Ottawa, Canada: Statistics Canada.

Vigo, M., & Brajnik, G. (2011). Automatic web accessibility metrics: Where we are and where we can go. *Interacting with Computers}, 23*(2), 137-155. doi: 10.1016/j.intcom.2011.01.001

Vigo, M., Brown, J., & Conway, V. (2013, May). Benchmarking web accessibility evaluation tools: measuring the harm of sole reliance on automated tests. In *Proceedings of the 10th International Cross-Disciplinary Conference on Web Accessibility* (W4A '13). ACM, New York, NY, USA, Article 1, 10 pages.

Wagner, N., Hassanein, K., & Head, M. (2010). Computer use by older adults: A multi-disciplinary review. *Computers in human behavior, 26*(5), 870-882.

Wallace, M., & Wray, A. (2009). *Critical reading and writing for postgraduates.* Thousand Oaks, CA: Sage.

Wandke, H., Sengpiel, M., & Soenksen, M. (2011). Myths about older people's use of information and communication technology. *Gerontology*, *58*(6), 564-570.

WebAIM. (2014). *WAVE: Web accessibility evaluation tool*. http://wave.webaim.org/

W3C. (2008). Web accessibility for older users: A literature review. *World Wide Web Consortium Working Draft 14 May 2008.* Retrieved from http://www.w3.org/TR/2008/WD-wai-age-literature-20080514/

W3C AGE. (2010). *Web Accessibility Initiative: Ageing Education and Harmonisation. Retrieved from http://www.w3org/WAI/older-users/Overview.php*

W3C WAI. (2005a). *Introduction to Web Accessibility*. Retrieved from http://www.w3.org/WAI/intro/accessibility.php

W3C WAI. (2005b). *Conformance evaluation of web sites for accessibility*. Retrieved from http://www.w3.org/WAI/eval/conformance.html

W3C WAI. (2008). *How WAI Develops Accessibility Guidelines through the W3C Process: Milestones and Opportunities to Contribute*. Retrieved from http://www.w3.org/WAI/intro/w3c-process

W3C WCAG 2.0. (2008). *Web Content Accessibility Guidelines*. Retrieved from http://www.w3.org/TR/WCAG/

White, P. (2009). *Developing research questions: A guide for social scientists*. Hampshire, United Kingdom: Palgrave Macmillan.

Yin, R.K. (2009). *Case study research: Design and methods* (4th ed.). Thousand Oaks, CA: Sage.

Zhang, L., Ayres, P., & Chan, K. (2011). Examining different types of collaborative learning in a complex computer-based environment: A cognitive load approach. *Computers in Human Behavior, 27*(1), 94. doi: 10.1016/j.chb.2010.03.038"

Appendix A - Pre-test questionnaire

In what year were you born?

What is your Gender?

What is your educational level? (high school, community college, university, graduate work)

What is your occupation now?

What was your previous occupation?

How would you rate your computer skills? 1-5, very poor to very good

How often do you use a computer?

What do you use a computer for?

On a scale of 1 to 7 with 1 being lowest and 7 being the highest, please rate the following as they apply to you: (for example, if you feel you have excellent hearing, you would rate this as 7).

Vision

Hearing

Memory

Movement/Mobility

Pain-Free

Appendix B – Test Tasks (Pilot Study Order)

Test tasks

Task 1:

Starting from the Ontario Ministry of Health website http://www.health.gov.on.ca/en/ participants will be asked to navigate to the proper form for renewing an OHIP health card.

Task 2:

Starting from the Ontario Ministry of Transportation website http://www.mto.gov.on.ca/english/ participants will be asked to navigate to the information page about renewing a driver's licence when a person is 80 years and older.

Task 3:

Starting from the Ontario Ministry of Labour website http://www.labour.gov.on.ca/english/ participants will be asked to find information about what a worker can do about unsafe conditions in the workplace.

Task 4:

Starting from the Ontario Ministry of Consumer Service website http://www.sse.gov.on.ca/mcs/en/Pages/default.aspx participants will be asked to find information about telemarketing scams.

Appendix C – Post-test Interview Questions

(for N=34 sample)

1. Can you tell me how you felt about this testing?
2. I noticed you hesitated/had difficulties at (a certain point/task). What was going on then?
3. How do you feel about government websites? What are your expectations of them?
4. Is there anything else you want to tell me about this testing? Your thoughts, feelings?

Appendix D - Semi-Structured Interview Questions

(for N=9 sample)

1. Describe your experiences with computers.
2. How often do you use the internet?
3. Where do you access it from?
4. What do you use it for?
5. What websites do you use the most?
6. Have you ever had any problems using the internet? What?
7. How did you learn to use the computer? Internet?
8. Do you have any physical limitations affecting your using the internet (e.g. reading glasses)?
9. How do you adapt to using websites? Is there anything you do to help yourself?
10. What suggestions would you have to make government websites easier to use?
11. Do you have anything else you wish to share about seniors using government websites?

Appendix E – Informed Consent Form, Pilot and Main Studies

Study name: Aging in Digital Worlds: How accessible are government websites to seniors?

Researcher: Kenneth H. Anderson
Doctoral Candidate in the Graduate Program in Education at York University
ken427@yorku.ca Phone: 905-736-5018

Purpose of the research: The central aim directing this study is to explore and determine how usable government websites are to seniors. The government websites that will be tested in the study are those that meet most technical accessibility standards. A primary objective is to gain insight into the self-reported experiences of seniors using government websites. The research will be conducted in the computer lab at (*To be named* Seniors Centre). The findings will be presented in my dissertation and may be presented through written and spoken means including conference papers and presentations, journal publications, books and book chapters, and/or electronic means (e.g. websites).

What you will be asked to do in the research: You will be asked to provide basic demographic information (age, gender, occupation, education, computer skill level), complete information search tasks (4) on government websites, complete a short questionnaire about your experiences with the search tasks, and answer interview questions about your internet experience (attached). You will be compensated for your time with a gift card in the amount of twenty-five dollars.

Risks and discomforts: You will be seated at a computer desk using a mouse, keyboard, and screen for a period of 30-45 minutes. You are asked to report any fatigue and/or discomfort so that a break can be taken during the session or the session ended. Since the Seniors Centre will be open at the time of the study, it is possible you may be recognized as a participant in the study.

Benefits of the research and benefits to you: The research will contribute to the literature on the design of accessible websites for seniors. It will also inform government policy on whether existing standards for website accessibility are sufficiently stringent to meet the needs of seniors.

Voluntary participation: Your participation in the study is completely voluntary and you may choose to stop participating at any time. Your decision not to volunteer will not influence the relationship you may have with the researchers or study staff or the nature of your relationship with York University either now, or in the future.

Withdrawal from the study: You can stop participating in the study at any time, for any reason, if you so decide. Your decision to stop participating, or to refuse to answer particular questions, will not affect your relationship with the researchers, York University, or any other group associated with this project. In the event you withdraw from the study, all associated data collected will be immediately destroyed wherever possible. If you decide to stop participating, you will still be eligible to receive the promised pay for agreeing to be in the project.

Confidentiality: Information collected during this research will be recorded in handwritten notes, audio recordings using a digital recorder, computer recordings of questionnaire responses, and computer screen saving recordings. The audio recordings and handwritten notes data will be stored on a computer that is password-protected, and in a locked filing cabinet, both accessible only to the researcher. The computer recordings will be stored on a password protected server in a locked room under the control of the Institute for Research on Learning Technologies (IRLT) at York University. All data will be destroyed two years after the study is complete. Confidentiality will be provided to the fullest extent possible by law.

Questions about the research? If you have any questions please contact either Kenneth H. Anderson or his supervisor Dr. Ron Owston at 905-736-2100 extension 66301. Alternatively you may contact the Faculty of Education Graduate Program Office at 416-736-5018. This research has been reviewed and approved by the Human Participants Review Sub-Committee, York University's Ethics Review Board and conforms to the standards of the Canadian Tri-Council Research Ethics guidelines. If you have any questions about this process, or about your rights as a participant in the study, you may contact the Senior Manager and Policy Advisor for the Office of Research Ethics, 5th Floor, York Research Tower, York University, telephone 416-736-5914 or e-mail ore@yorku.ca

Legal rights and signatures:

I, , consent to participate in the research project "Aging in digital worlds: How accessible are government websites to seniors?" conducted by Kenneth H. Anderson . I have understood the nature of this project and wish to participate. I am not waiving any of my legal rights by signing this form. My signature below indicates my consent.

Signature _____ **Date** _____
Participant

Signature _____ **Date** _____
Principal Investigator

Additional consent:

□ I consent to the audio recording of my voice during the study.

□ I consent to the computer recording of my questionnaire responses and website navigation during the study.

Signature _____ **Date** _____
Participant

Signature _____ **Date** _____
Principal Investigator

Appendix F – Informed Consent Form, Semi-Structured Interview

Study name: Aging in Digital Worlds: How accessible are government websites to seniors?

Researcher: Kenneth H. Anderson
Doctoral Candidate in the Graduate Program in Education at York University
ken427@yorku.ca Phone: 905-736-5018

Purpose of the research: The central aim directing this study is to explore and determine how usable government websites are to seniors. The government websites that will be tested in the study are those that meet most technical accessibility standards. A primary objective is to gain insight into the self-reported experiences of seniors using government websites. The research will be conducted in the computer lab at (*To be named* Seniors Centre). The findings will be presented in my dissertation and may be presented through written and spoken means including conference papers and presentations, journal publications, books and book chapters, and/or electronic means (e.g. websites).

What you will be asked to do in the research: You will be asked to answer approximately ten interview questions about your internet experience, and engage in a conversation with the researcher around these questions. This may take about 30 minutes to complete. You will be compensated for your time with a gift card in the amount of twenty-five dollars.

Risks and discomforts: You are asked to report any fatigue and/or discomfort so that a break can be taken during the session or the session ended. Since the Seniors Centre will be open at the time of the study, it is possible you may be recognized as a participant in the study.

Benefits of the research and benefits to you: The research will contribute to the literature on the design of accessible websites for seniors. It will also inform government policy on whether existing standards for website accessibility are sufficiently stringent to meet the needs of seniors.

Voluntary participation: Your participation in the study is completely voluntary and you may choose to stop participating at any time. Your decision not to volunteer will not influence the relationship you may have with the researchers or study staff or the nature of your relationship with York University either now, or in the future.

Withdrawal from the study: You can stop participating in the study at any time, for any reason, if you so decide. Your decision to stop participating, or to refuse to answer particular questions, will not affect your relationship with the researchers, York University, or any other group associated with this project. In the event you withdraw from the study, all associated data collected will be immediately destroyed wherever possible. If you decide to stop participating, you will still be eligible to receive the promised pay for agreeing to be in the project.

Confidentiality: Information collected during this research will be recorded in handwritten notes, and audio recordings using a digital recorder. The audio recordings and handwritten notes data will be stored on a computer that is password-protected, and in a locked filing cabinet, both

accessible only to the researcher. All data will be destroyed two years after the study is complete. Confidentiality will be provided to the fullest extent possible by law.

Questions about the research? If you have any questions please contact either Kenneth H. Anderson or his supervisor Dr. Ron Owston at 905-736-2100 extension 66301. Alternatively you may contact the Faculty of Education Graduate Program Office at 416-736-5018. This research has been reviewed and approved by the Human Participants Review Sub-Committee, York University's Ethics Review Board and conforms to the standards of the Canadian Tri-Council Research Ethics guidelines. If you have any questions about this process, or about your rights as a participant in the study, you may contact the Senior Manager and Policy Advisor for the Office of Research Ethics, 5th Floor, York Research Tower, York University, telephone 416-736-5914 or e-mail ore@yorku.ca

Legal rights and signatures:

I, , consent to participate in the research project "Aging in digital worlds: How accessible are government websites to seniors?" conducted by Kenneth H. Anderson . I have understood the nature of this project and wish to participate. I am not waiving any of my legal rights by signing this form. My signature below indicates my consent.

Signature_____ **Date**_____
Participant

Signature_____ **Date**_____
Principal Investigator

Additional consent:

□ I consent to the audio recording of my voice during the study.

Signature_____ **Date**_____
Participant

Signature_____ **Date**_____
Principal Investigator

Appendix G –Timetable

Mar-Aug 2013	-draft(s) dissertation proposal submitted, comprehensive exam scheduled, study site selected, approvals obtained
Sept 2013	-comprehensive exam completed, ethics submission
Oct-Dec, 2013	-ethics approvals obtained, site approvals obtained, pilot study completed
Jan 2014 – June 2014	-main study data collection
Jul 2014 – Oct 2014	-completion of data analysis
Nov 2014 – Apr 2015	-completion of draft dissertation writing
May 2015	-submission of draft dissertation to supervisor
June 2015	-submission of draft dissertation to committee

Appendix H – Initial Qualitative Codes

1. Adaptations made
2. Previous experience
3. Ipad or smartphone usage
4. Persistence
5. I'm missing something
6. Reading the web pages
7. Uses search function
8. Seeking Help
9. Making Notes
10. Feelings
11. Feeling Stupid
12. Frustration
13. I think it's me
14. I'm beginning to feel smart
15. Proud of myself
16. Technical Issues
17. Colour Contrast on Health Site
18. Slow Computer
19. Problems with the websites
20. Confusion with question
21. Not enough clarity
22. Navigation issues
23. Jumping around
24. Lost
25. Too much information
26. Perceptions expressed
27. Evaluations given
28. Competence perceived
29. Motivation to use websites
30. Perceived usability
31. Websites are not easy
32. Websites are easy
33. Websites are good
34. Websites are bad
35. Seniors and government websites
36. Suggestions for government websites