E-Learning in Higher Education: Designing for Diversity

Kari L. Kumar

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

September, 2013

Abstract

This research was conducted to compare methods of e-learning accessibility evaluation that may be applied in a higher education context. Results of "objective" accessibility evaluation of e-learning technologies using automated tools were compared to results of "subjective" accessibility evaluation with student participants. It was found that objective and subjective accessibility evaluation of e-learning technologies both yield useful, albeit different, information. To further explore subjective accessibility evaluation, results and student perceptions were compared following moderated and unmoderated testing sessions. Neither the efficiency of completing tasks in a sample online course nor the number of accessibility problems detected were deemed significantly affected by the format of the testing session. However, most students preferred to participate in an unmoderated testing session where they felt less self-conscious and as though they could interact more naturally with the technology. Findings from this study point to the importance of considering not only objective accessibility evaluation and accessibility guideline conformance as measures of the accessibility of e-learning technologies, but also the subjective experiences of students as they engage with the technologies. There is also value in taking a holistic approach towards evaluating e-learning accessibility by considering the accessibility of learning outcomes (factoring in the learning context to the evaluation) in addition to the accessibility of individual e-learning technologies. Because accessibility is a variable that is important to all students, and not just students with disabilities, it is critical that institutions of higher education work with a variety of

stakeholders to determine not only how best to evaluate e-learning accessibility, but also how to ensure that the results of accessibility evaluation are widely disseminated in a manner that is likely to have a broad impact on enhancing e-learning accessibility for diverse student populations.

Dedication

For Divya and Ayush, each equally my pride and joy.

Acknowledgements

I would like to express deep gratitude to my supervisor, Dr. Ron Owston, for the tremendous support and mentorship that he has provided me throughout the completion of this work. I would also like to thank the members of my supervisory committee, Dr. Neita Israelite and Dr. Melanie Baljko, for sharing their expertise and helping to shape both the project design and presentation of this dissertation. I would also like to acknowledge the technical assistance provided by Jacky Siu and Sydney Collins during the data collection stage of this project. Last, but certainly not least, thank you to the students who took part in this study and generously shared their time and ideas.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1: Introduction

The overarching goal of this research project was to garner insight towards understanding how to create post-secondary learning environments that are as inclusive as possible for diverse populations of learners. In this context, an inclusive learning environment has been deemed one in which all students have an equitable opportunity to succeed. One key ingredient of an inclusive learning environment is accessibility, which may be understood as the extent to which the learning environment meets the needs of learners by exhibiting properties of flexibility and adaptability. This notion of the value of flexibility and adaptability as essential ingredients within an accessible classroom supports an examination of learning technologies. This is because there is a wide range of learning technologies available and e-learning may therefore encompass many different possibilities. Could e-learning, with much potential for built-in flexibility and adaptability, be an emerging platform for more fully supporting the diverse needs and preferences of our students, by helping to reduce mismatches between the learners and the learning environment and thus enhance accessibility?

In order to explore the extent to which e-learning may enhance the inclusiveness of higher education, it is necessary to consider the degree to which e-learning environments – and the learning technologies used therein – are accessible. In other words, are the variety of learner needs and preferences (e.g., related to how they may access, interact with, and/or understand learning materials) met by the e-learning environment that they encounter? To begin to address this broad question, one essential avenue to explore is the characterization of data obtained from different methodologies that may be used to evaluate the accessibility of e-learning technologies. The necessity for exploration in the area of methods of e-learning accessibility evaluation has informed the research objectives for this study, as described in the following section.

1.1 Research Objectives

This research was conducted with the objective of evaluating methods to assess the accessibility of learning technologies used in higher education. The specific research objectives of this project were to:

1. Determine the extent to which objective measures of the accessibility of e-learning technologies are able to predict the subjective accessibility experience of students, and to
2. Determine whether data obtained from moderated and unmoderated e-learning accessibility testing are different and, if so, how and why they differ.

In this dissertation, "objective" testing refers to automated accessibility testing of technology, while "subjective" testing refers to testing of technology by anticipated users (i.e., students).

1.2 Significance and Rationale of the Study

Due to the increasing use of technology in face-to-face classes and the rise of blended and online learning in higher education, there is a pressing need to develop tools and strategies to examine the accessibility of e-learning environments. Decades of research comparing student learning outcomes in the presence and absence of learning technologies has shown that technology may enhance learning at all levels of education, from K-12 to postsecondary education (Tamim, Bernard, Borokhovski, Abrami, & Schmid, 2011). However, the full benefits of e-learning will not be realized if e-learning platforms and materials are not sufficiently flexible such that they may be used and/or understood by all students. In other words, accessibility is a key factor that will influence the degree of benefit that students may obtain from e-learning.

All students may encounter barriers to learning to varying degrees and in various contexts, and therefore all students stand to benefit from an overall increase in e-learning accessibility. It is also noteworthy that participation rates of students with disabilities[1] in higher education are on an upward trend. In 1993, Statistics Canada estimated that approximately 7% of post-secondary students in Canada have one or more disabilities (as cited by Hubka & Killean, 1999), and a more recent report by the National Educational Association of Disabled Students (NEADS) of Canada has indicated that up to 10% of students enrolled in college and university across the country may have a disability (NEADS, 2005). This correlates with data from the United States, with recent estimates suggesting that 11% to 14.6% of the post-secondary student population may have a disability (Higher Education Research Institute, 2011; U.S. Government Accountability Office, 2009). Students with disabilities potentially have much to gain from e-learning

[1] In Canada, the use of person-first language (e.g., students with disabilities) predominates. However, some feel that this language lacks acknowledgement that people may be disabled by society (for example, see Titchkosky, 2001), and it is therefore also acceptable to use disability-first language (e.g., disabled students). Both forms of language are used within this dissertation.

due to the possibility of inherently flexible properties of learning technologies, yet may also be disadvantaged if inaccessible e-learning technologies enhance barriers currently experienced in face-to-face learning environments and/or create new barriers to learning (Cook & Gladhart, 2002). This observation, coupled with the wide range in needs and preferences that students with disabilities collectively represent, makes students with disabilities an ideal population from which to solicit insight into how to design e-learning environments with student diversity in mind.

Legislative compliance is also an emergent variable to consider with respect to e-learning accessibility. Several countries have adopted policies or legislation to promote or enforce inclusive technology-enhanced education. In Canada, such matters fall under provincial jurisdiction. The Integrated Accessibility Standards of the Accessibility for Ontarians with Disabilities ACT (AODA Integrated Accessibility Standards, 2011) mandate that all websites of post-secondary institutions in the province (including posted content and applications that appear therein) adhere to a Web accessibility conformance schedule that is based on internationally-recognized benchmarks for Web accessibility, namely, the WCAG 2.0 (W3C, 2008). Specifically, all new websites created after June 2011 must become WCAG 2.0 conformant by January 2014, and all websites must become conformant by January 2021. While this is encouraging, there are only preliminary interpretations of data appearing in the literature regarding strengths and limitations of the WCAG 2.0 with respect to informing accessible Web design, serving as reliably testable conformance standards, and meeting the needs of Web users from all disability groups (for example, refer to Alonso, Fuertes, González, & Martinez, 2010;

Brajnik, 2009; Seeman, 2006a, 2006b). The general consensus to-date is that the guidelines are very useful, though conformance may not necessarily equate with accessibility. Without identifying how subjective feedback from students regarding e-learning accessibility compares to the results of accessibility guideline conformance testing (Research Objective 1), it is unclear as to how effective extant legal measures can be if accessibility is defined primarily in relation to guideline conformance.

There is a dearth of peer-reviewed studies in the literature that report results of accessibility testing of new learning technologies by soliciting feedback from student users. There are reports of the use of traditional usability testing techniques for evaluation of academic library digital content (for examples, refer to Bury & Oud, 2004; Denton & Coysh, 2011; Dermody & Majekodunmi, 2010) and there is emerging interest in applying usability testing methods to accessibility testing of a wider array of educational technologies (see Power, Petrie, Sakharov, & Swallow, 2010 for example). There are many variations of usability testing methods, and is it therefore important to identify which testing method(s) may be best suited to be administered in modified (accessibility-focused) form for use with student participants, including students with disabilities.

Traditional usability testing takes place in a formal testing laboratory which requires users to visit the laboratory and to work with equipment that has been provided for them (Nielsen, 1993). There is currently emerging interest in unmoderated usability testing such as remote testing, which may take place over the Internet, where users may participate when, where, and how (i.e., using their own computer and software configurations) they prefer (Bolt & Tulathimutte, 2010; Hartson, Castillo, Kelso, &

Neale, 1996), and where there is not real-time moderation from the researcher(s). It is therefore highly relevant to compare both the richness of data and student experiences with respect to traditional moderated and unmoderated accessibility testing based on usability testing methods (Research Objective 2).

This study took place at York University, where e-learning accessibility is highly relevant. York University is a large Toronto area university with more than 50 000 undergraduate and graduate students (http://www.yorku.ca/web/about_yorku/) and, in 2009, more than 12 000 York University students were enrolled in 141 fully or partially online courses (Owston, 2010). The university has expressed a commitment towards "significantly expanding online delivery of courses and programs as part of its efforts to enhance learning through the use of technology" (Monahan, 2010). Since York University is a public university in the province of Ontario, its e-learning content is subject to the Web content accessibility conformance schedule that appears in the AODA Integrated Accessibility Standards and which is based on the WCAG 2.0. As such, the university has strong moral and legal reasons for ensuring that e-learning is accessible.

CHAPTER 2: Literature Review

Abstract

E-learning is becoming ubiquitous in higher education, and an important new direction in e-learning research is to identify how to maximize the positive impact that learning with technology may have on student learning outcomes. Because students with disabilities represent an increasing proportion of students in post-secondary education and because they collectively have a wide range of learning needs and preferences, it is useful to include their perspectives when identifying how accessible typical e-learning environments may be for diverse student populations. A review of the literature reveals that there is indeed much potential of learning technologies to enhance accessibility of learning for all students but perhaps just as many frequently-encountered challenges associated with inaccessible e-learning environments, particularly for students with disabilities. An overview of basic steps that may be taken by instructors to enhance accessibility of e-learning infrastructure is presented, followed by an in-depth discussion of Web accessibility guidelines that are currently legally enforceable with respect to post-secondary e-learning in many jurisdictions, including Ontario. Because accessibility is a somewhat subjective property, tools and methods for conducting e-learning accessibility testing with students are needed to complement objective guideline conformance testing. An exploration of the intersection between accessibility and usability (described by Nielsen, 1993, as encompassing learnability, efficiency, memorability, low error rate, and user satisfaction – refer to section *2.6.1 Relationship between accessibility and usability*

for a more in-depth discussion) reveals that well-established usability testing methods may be applied to e-learning accessibility testing.

2.1 E-Learning in Higher Education

Use of the term "e-learning" is becoming ubiquitous in educational contexts. The "e" in e-learning represents "electronic" and this term refers to learning that takes place when some or all learning material is accessed electronically. This involves the use of computers or other electronic devices (such as smart phones or iPads) to access content stored on CDs, DVDs, or the Internet. Learning Management Systems (LMSs) are software packages designed for the delivery of online instruction and are frequently used by teachers to post course materials that students may access over the Internet (White & Larusson, 2010).

Given the widespread use of the Internet and the World Wide Web[2] in higher education, the term "e-learning" often brings to mind online learning. However, e-learning is a broad term that encompasses various forms of course delivery including not only fully online courses, but blended courses and technology-enhanced face-to-face

---

[2] The terms "Internet" and "World Wide Web" (or simply, "the Web") are often used interchangeably in education literature, though there is a distinction between the two. The Internet is the actual network connecting computers, while the World Wide Web refers to the use of Web browsers to navigate through online information via hyperlinks (Ellcessor, 2010). In other words, using the Web is one way of accessing information available on the Internet.

(FTF) courses as well. Courses delivered solely via the Internet are referred to as online courses, courses in which there is a combination of FTF and online delivery of content are referred to as hybrid or blended courses, and courses in which all FTF time is retained but which include the use of technology may be referred to as technology-enhanced or Web-enhanced courses (Bates, 2005).

The prevalence of e-learning in its various forms is increasing in higher education. There are several factors that have contributed to this rise in e-learning, including: the potential for learning technologies to enhance learning outcomes and learner satisfaction, the potential for online learning to assist institutions in meeting enrollment pressures, and the potential of e-learning (particularly online learning) to expand opportunities for students traditionally under-represented in higher education (Abrami et al., 2006; Advisory Committee for Online Learning, 2001; Allen & Seaman, 2010; Canadian Council on Learning, 2009). Allen and Seaman's large scale study on online learning in U.S. higher education in 2010 (Allen & Seaman, 2011) reported that rates of enrollment in higher education are increasing but enrollment in online courses is increasing at a faster rate. They have reported a 10% increase in the number of students taking at least one online course in 2010 compared to 2009, while the overall higher education student population has been increasing at an annual rate of about 2% (Allen & Seaman, 2011).

E-learning is also prevalent in Canada. For example, Athabasca University, a distance education university in Canada, reports enrollment exceeding 38 000 students. There are also several Canadian consortia offering online courses including Contact North (with 17 partner colleges and universities), Ontario Learn (with 22 partner colleges), and Canadian Virtual University (with 10 partner universities). There is interest in further expanding e-learning opportunities in post-secondary education in Canada (Advisory Committee for Online Learning, 2001; Canadian Council on Learning, 2009). For example, in the 2010 Throne Speech the Ontario provincial government announced plans for an Ontario Online Institute – a new consortium of Ontario universities intended to increase accessibility to Ontario universities and to attract foreign students (Office of the Premier Dalton McGuinty, 2010).

Given the widespread availability of post-secondary e-learning, it is not surprising that reports of technology use by students and faculty also reveal positive trends. For example, the recent EDUCAUSE Core Data Service survey of information technology environments and practices in higher education reported that more than 90% of the 875 institutions that participated in the survey use LMSs (e.g., Blackboard, Moodle, or WebCT) and that faculty use of LMSs is on the rise (Arroway, Davenport, & Xu, 2010). This correlates with data from the 2010 ECAR study of undergraduate students and information technology, which reports that more than 90% of students have used an LMS (S. D. Smith, Saloway, & Caruso, 2010). Moreover, recent U.S. data also suggests that post-secondary students, in general, are more likely to be Internet users and to have a laptop compared to all adults and non-students aged 18-24 (A. Smith, Rainie, & Zickuhr, 2011).

2.2 Past Trends and New Directions in E-Learning Research

As the use of learning technologies has increased in all educational levels, the literature has been inundated with studies examining the impact of technology on learning. Particularly popular are comparative studies in which student achievement in the presence and absence of learning technologies has been compared. Hundreds of such studies have been conducted over the last 25 years, since computer use became fairly widespread in classrooms in the mid 1980's. For example, Bernard and colleagues (2004) identified 232 primary studies published between 1985 and 2002 that compared modern distance education (incorporating e-learning elements) with traditional classroom instruction.

As the number of primary comparative studies has grown, so too has the number of meta-analyses on the topic. Meta-analyses are a particularly useful tool in this field due to the wide variety in methods and sample sizes in the primary e-learning literature. Bernard, Abrami, Lou, and Borokhovski (2004) speculate that the wide disparity in research methodologies of studies pertaining to distance education, for example, may reflect unique challenges in the field. Challenges include the difficulty of designing true experimental studies with students randomly assigned to different treatment groups, difficulties in recruiting online learners to participate in studies, and loss of participants during the course of a study (due to relatively high attrition rates in distance education). Herein lies the value of meta-analysis: Meta-analyses employ statistical methods to integrate findings and draw generalized conclusions from many primary studies that address the same problem but do not necessarily employ the same experimental design (Glass, McGaw, & Smith, 1981). A particularly useful metric used in meta-analyses is the effect size – the difference between two treatments (e.g., learning FTF or online) divided by the standard deviation (Glass, et al., 1981). The effect size is an indication of the size of an effect a particular treatment has and, in contrast to significance testing, is not subject to a large influence from sample size and allows for direct comparison of the overall findings from different studies.

With a large collection of comparative primary studies and corresponding meta-analyses available in the literature, a second-order meta-analysis has recently been published that summarizes the trends from this research area. Tamim and colleagues (2011) included 25 comparative meta-analyses representing 1055 primary comparative studies in their second-order meta-analysis. By retrieving effect sizes from individual meta-analyses (including data from studies on K-12 and postsecondary classrooms) and conducting additional statistical analyses to further standardize the data from different studies, the authors report an average effect size of 0.33. This effect size may be expressed in percentile points as follows: "the average student in a classroom where technology is used will perform 12 percentile points higher than the average student in the traditional setting that does not use technology to enhance the learning process" (Tamim, et al., 2011, p. 17). Moderator analyses revealed that technology that was used to support instruction and technology used in K-12 settings led to higher average effect sizes than technology used to provide direct instruction and that used in post-secondary settings, respectively. The conclusion from this paper and the large number of primary studies that it represents suggests that:

a) Learning technology (and thus e-learning) supports and may even enhance student achievement, and

b) It is useful to move beyond the traditional comparative study to focus on new areas of research relevant to e-learning.

A key question to explore now is not whether learning technologies may be beneficial but rather, how to maximize their potential to enhance learning. While the moderate positive effect of learning technology on student achievement reported by Tamim et al. (2011) is encouraging, the results of many years of implementation and study about e-learning begs the question: what can be done to increase the positive effect of technology on learning? Given the increased presence of learning technologies in the classroom, it stands to reason that the "traditional classroom" included as a control in many comparative e-learning studies of the past may itself evolve as technology becomes more ubiquitous in education. As such, studies aimed at examining student outcomes in modern day classrooms may lend insight about strategies to enhance learning in technology-enhanced face-to-face (FTF) and blended classes. Moreover, new studies that examine the impact of different treatments within online classes (rather than comparing online vs. FTF classes) will further an understanding of how to enhance online learning based on attributes that are unique to online learning environments.

In the context of higher education, another avenue of research to explore is related to the motivation for increased prevalence of e-learning. In addition to potential administrative benefits, several reports recommend offering more e-learning options to open up the doors of higher education to students who are traditionally excluded or present in relatively low numbers. E-learning (online learning in particular) is often touted as a flexible alternative to traditional FTF post-secondary instruction that is more practical and accessible to working students, students with dependents, rural students, Aboriginal students, and students with disabilities (Advisory Committee for Online Learning, 2001; Canadian Council on Learning, 2009). For example, the flexibility with respect to when and where online learning takes place could be beneficial for working parents who may find it necessary to focus on studies in the evening, and for rural and Aboriginal students who prefer not to move away from their community and relocate to an urban environment to pursue a degree. Indeed, flexibility is an attribute that may be enjoyed by all students, and the benefits of e-learning may increase for all students if e-learning environments are designed to be flexible with student diversity in mind. It is therefore useful to explore how e-learning may increase inclusion in higher education, by increasing flexibility and thus accessibility.

2.3 E-learning Accessibility and Post-Secondary Students

When considering the accessibility of learning technologies from the student perspective, it is useful to include the vantage point of students with disabilities. This is because students with disabilities collectively represent a wide spectrum of learning needs and preferences, and students without disabilities may also identify with similar needs and preferences in certain contexts – it could be said that accessibility from the vantage point of students with disabilities may be the ultimate acid test of accessibility. As such, this section will begin with a brief discussion of the nature of various disabilities that students in higher education may identify with and some of the ways in which the disabilities may impact a student's learning, prior to describing benefits and drawbacks of e-learning with respect to the accessibility of higher education for diverse student populations.

2.3.1 Students with disabilities in higher education. In order to be eligible for disability services in post-secondary institutes (e.g., accommodations in which the method of instruction or assessment may be altered in response to a student's needs), students must typically undergo some form of formal assessment and provide documentation of their disability or disabilities. For example, this may include a student providing medical documentation to a campus disability services office to verify a sensory impairment in order to be eligible to receive alternate format course materials. This method of labelling a student with a disability resulting from impairment is aligned with the medical model of disability which emphasizes individual impairment as the cause of disability (Marks, 1997). While this model is often the basis of determining whether or not a post-secondary student is classified as having a disability of one or more categories, most contemporary discussions of disability take into account the role of societal factors in contributing to disability. For example, the WHO characterizes disability as follows:

> Disabilities is an umbrella term, covering impairments, activity limitations, and participation restrictions. An impairment is a problem in body function or structure; an activity limitation is a difficulty encountered by an individual in executing a task or action; while a participation restriction is a problem experienced by an individual in involvement in life situations. Thus disability is a

complex phenomenon, reflecting an interaction between features of a person's body and features of the society in which he or she lives. (WHO, 2011)

The current WHO characterization of disability suggests that disability is not necessarily solely attributed to impairment, and that disability may be contextualized. For example, students who are print-disabled as a result of an impairment that affects their ability to read text may not identify as being disabled when provided with audio in lieu of text course materials. Additional discussion of the social construction of disability in e-learning environments appears in section *3.1 Defining Digital Disability and E-Learning Accessibility*, of Chapter 3.

Defining categories of disability is not as simple a task as one might expect, as definitions of disability categories vary by country and jurisdiction. For example, learning disabilities are classified as types of intellectual disabilities by United Kingdom government departments, however, learning disabilities and intellectual disabilities are typically considered distinct outside of the United Kingdom (Kennedy, Evans, & Thomas, 2011). Differences in how disability categories are defined are also found within a given country. For example, definitions of disability categories are not consistent across the provinces in Canada (Winzer, 2008). The following discussion is an effort to describe definitions of disability categories that predominate in Canadian educational contexts by drawing primarily on Canadian literature and examples.

2.3.1.1 Learning disabilities. The term "learning disability" is particularly difficult to define due to the highly heterogeneous nature of learning disabilities. Many definitional variations exist between provinces in Canada, and Winzer (2008) suggests that variability in identification/definition of learning disabilities may even exist at the level of individual teachers within a given province. According to the Learning Disabilities Association of Canada (LDAC):

> Learning Disabilities refer to a number of disorders which may affect the acquisition, organization, retention, understanding or use of verbal or nonverbal information. These disorders affect learning in individuals who otherwise demonstrate at least average abilities essential for thinking and/or reasoning. As such, learning disabilities are distinct from global intellectual deficiency. (LDAC, 2002)

Among those disabilities that may be included within the umbrella term of learning disability are dyslexia, dysgraphia, dyscalculia, and attention deficit disorder (Bohman & Anderson, 2005). Bohman and Anderson (2005) suggest that functional categorization of learning disabilities is a valuable approach to understand the impact that various learning disabilities may have, and the LDAC provides a general functional description of learning disabilities that is helpful in this regard. According to the LDAC, learning disabilities may impact skills related to oral language (including listening, speaking, and/or understanding), reading, written language, and mathematics, in addition to other skills including organization and social perception (LDAC, 2002).

The terms "learning disability" and "cognitive disability" are generally considered to be overlapping as manifestations may be very similar. However, a distinction is sometimes made between these terms. Learning disabilities may be considered disabilities that a person is born with (e.g., dyslexia), while cognitive disabilities may be considered disabilities that are acquired later in life (e.g., as a result of an acquired brain injury) (Kennedy, et al., 2011). An inspection of disability services resources of several Canadian universities suggests that learning disabilities and cognitive disabilities are not clearly distinguished in terms of service provision to students, and that the term "learning disabilities" predominates. As such, the term "learning disability" will be preferentially used throughout the remainder of this dissertation.

Students with learning disabilities represent a large and growing proportion of students with disabilities in higher education. Statistics Canada has found that learning disabilities are the most prevalent type of disability amongst Canadian children (as cited by LDAC, n.d.), with an estimate that 59.8% of disabled children in the country identify with learning disabilities. Likewise, the Toronto District School Board of Toronto, Canada has reported that students with learning disabilities comprised 42.1% of all reported students with exceptionalities in 2005 (Brown, 2008) and Wolanin and Steele (2004) report that the proportion of students with learning disabilities in the United States may range from 46% to 61%. These estimates correlate with data from higher education. For example, a recent study by Fichten and colleagues (2009) found that, of 223 post-secondary students with disabilities surveyed across Canada, 41% identified as persons with learning disabilities.

2.3.1.2 Mental health disabilities. The Canadian Mental Health Association (CMHA) indicates that mental illnesses may also be referred to as psychiatric disorders, and that a wide variety of conditions including mood disorders, anxiety disorders, eating disorders, and schizophrenia fall under the umbrella of mental illness (CMHA, n.d.-b).

Post-secondary students may identify with a variety of mental illnesses. Indeed, Mental Health Disability Services (MHDS) at York University in Toronto, ON includes a similar list of examples of mental health disabilities (York University, n.d.).

Students with mental illness frequently face a variety of barriers that may act as obstacles to their learning. These barriers may be related to the nature of the illness itself, the effects of medication, and/or the physical or social climate of the learning environment. For example, the illness itself or medication taken to treat the illness may slow cognitive processing (Eudaly, 2003). In a survey of 387 adults with mental illness following completion of college preparatory classes, students reported barriers to completion of the program that included difficulty concentrating in class, attending class, accepting criticism, taking notes, joining class discussions, and getting along with other students (Mowbray, Bybee, & Collins, 2001). Similarly, a study conducted by Weiner and Weiner (1996) in which 24 post-secondary students with mental illness were surveyed revealed that barriers to learning included problems focusing, short attention span, lack of self-esteem, and difficulty trusting others. Challenges faced can also arise from the teaching style and approach, as college students with mental illness in a study conducted by Blacklock, Benson, and Johnson (2003) commented that an unwelcoming classroom climate hindered their success. Stereotypes and stigma associated with mental illness have also been reported as problematic by students with mental illness (Blacklock, et al., 2003; Eudaly, 2003; Weiner & Weiner, 1996).

It is particularly difficult to estimate the number of students in higher education with mental health illness. The Canadian Mental Health Association suggests that 20% of all Canadians will experience mental illness firsthand (CMHA, n.d.-a). A small study by Offer and Pollack (1987) found that approximately 20% of the 60 high school graduates that they studied exhibited mental illness, suggesting that the proportion of post-secondary students with mental illness may mirror that of the general public. This correlates with Fichten and colleagues' recent Canadian study in which 17% of the 223 student respondents with disabilities self-reported as having a psychological or psychiatric illness (Fichten, et al., 2009). However, difficulty estimating the number of post-secondary students with mental illness arises for several reasons. One reason is that diagnosis often doesn't occur until adulthood (Beiser, Erickson, Fleming, & Iacono, 1993; Kessler et al., 2007) and students may therefore not know that they have a mental health illness when they begin college or university. A second reason is the perceived stigma associated with mental illness (Blacklock, et al., 2003; Eudaly, 2003; Weiner & Weiner, 1996) that may prevent students who have been diagnosed from disclosing their illness. However, an increase in the number of students with mental health illness in higher education has been reported (Eudaly, 2003; Pledge, Lapan, Heppner, Kivlighan, & Roehlke, 1998; Sharpe, Bruininks, Blacklock, Benson, & Johnson, 2004). Indeed, many post-secondary institutions have reported substantial increases in numbers of students disclosing mental illness (Fichten, Jorgensen, Havel, & Barile, 2006; Sharpe, et al., 2004), with some estimates that rates of growth of students with mental illness seeking support on campus may have reached 50% (Grabinger, 2010).

2.3.1.3 Sensory disabilities. Sensory disabilities are those related to auditory and visual function. The term "hearing impairment" is a general term that refers to some degree of hearing disability (Winzer, 2008). Students with hearing impairment may consider themselves to be hard-of-hearing (e.g., a student who uses a hearing aid) or deaf (Hutchinson, 2010). Moreover, students may consider themselves to be culturally Deaf – heretofore referred to as *Deaf* – if they communicate via American Sign Language regardless of the degree of hearing loss. Similarly, different terminology may be used to describe different degrees of impaired visual function, with the term "visual impairment" generally denoting partial vision loss, and "blind" referring to more severe or complete vision loss (Hutchinson, 2010, p. 27). In an educational context, students are deemed to be students with visual impairments if their visual impairment requires adaptations to be made to their learning environment (Winzer, 2008).

Post-secondary students with sensory disabilities may experience a variety of barriers to learning. Students with visual impairments or who are blind may not be able to receive information that is presented visually in the class (Lewin-Jones & Hodgson, 2004), particularly if the lighting is low (Bishop & Rhind, 2011). Alternate formats of course materials can be helpful, though students report a delay in receiving alternate format materials (Bishop & Rhind, 2011). Moreover, students with visual impairment may find that they require a longer time to process written information (Owen Hutchison, Atkinson, & Orpwood, 1998) and that a large volume of reading can lead to headache, body ache, and fatigue (Bishop & Rhind, 2011; Owen Hutchison, et al., 1998). Assistive technologies such as text-to-speech software can help to alleviate these challenges, though students have expressed frustration at the expense and learning curve related to assistive technologies (Owen Hutchison, et al., 1998) and, when used in the classroom alongside other students, the assistive technologies may be distracting to other students who are working quietly (Lewin-Jones & Hodgson, 2004). In addition to challenges receiving and processing information, students with visual impairment may experience challenges physically navigating through the campus. Inadequate signage or lighting, difficulty reading maps, and difficulty adapting to lighting changes in different campus locations can all be challenging (Bishop & Rhind, 2011; Owen Hutchison, et al., 1998).

Students with hearing impairment may have difficulty receiving information that is presented orally in the class. Speechreading can be helpful, though is challenging when teachers move around a lot, obstruct their face when talking, speak in a dimly lit location, or when more than one person is speaking at a time (Hyde et al., 2009; Waterfield & West, 2002). Moreover, it can be particularly difficult for a student with a hearing impairment to listen, learn, and take notes simultaneously (Hyde, et al., 2009). As a result, students with hearing impairment often need help to receive information presented in the class, such as from an interpreter, note-taker, or captioning (Hyde, et al., 2009; Lang, 2002). However, the usefulness of this information is affected by the accuracy or skill of the third party, and students have reported that they do not receive all information that is presented in the class even when this assistance is provided (Hyde, et al., 2009; Marschark, Sapere, Convertino, & Seewagen, 2005; Napier & Barker, 2004).

As observed with other categories of disability, enrollment of students with sensory disabilities is on an upward trend. For example, following a review of statistics from universities in the United States and Australia, Lang (2002) found that the number of students with hearing impairment enrollment in higher education has been increasing since the 1980's. Recent data from Fichten and colleagues' Canadian study (2009) found that 13% of the 223 student participants with disabilities in their study self-reported as having a visual impairment (low vision or blindness), and 13% self-reported as being *Deaf* or hard-of-hearing.

2.3.1.4 Physical disabilities. Physical disabilities may be described as "a range of conditions restricting physical movement or motor abilities as a result of nervous system impairment, musculoskeletal conditions, or chronic medical disorders" (Hutchinson, 2010, p. 27). From this definition, it is evident that physical disabilities may occur as a result of a very wide range of conditions (e.g., paralysis, epilepsy, migraines, cancer), some of which may result in episodic impairment while other conditions may result in chronic impairment. Moreover, the impact of a physical disability on the learning of students may vary tremendously depending on the nature and the severity of the impairment experienced by any individual student at any point in time. Fichten and colleagues (2009) have used the subcategories mobility impairment/wheelchair user, difficulty using hand and/or arms, and health/medically-related impairment to describe physical disabilities, and found that 23, 12, and 16%, respectively, of the student respondents in their Canadian study self-reported with these types of physical disabilities.

2.3.1.5 Key observations. This overview of students with disabilities in higher education has revealed that there are students with many different types of disabilities enrolled in higher education, and that rates of enrollment of students with disabilities across the various disability categories are rising. While there are barriers that may be unique to or more commonly experienced by students with certain types of disabilities, there are also many commonalities amongst the barriers that students with disabilities may face in a post-secondary setting. For example, students with learning disabilities may have a strong preference for either written or auditory learning materials, as may students with sensory disabilities. Students with all categories of disabilities may find it difficult to pay attention or take notes in class, albeit for different underlying reasons. Students with mental health or physical disabilities may find that their illness is episodic, and that barriers are present or more pronounced at different times of the academic year.

Another key observation that arises from this review of the literature on students with disabilities in higher education is that students who do not identify as persons with disabilities can identify with many of the same barriers in certain contexts. For example, a student who works to support him/herself while attending college or university may feel fatigued if attending class after working, and may experience difficulties receiving and processing information that may be similar to students with disabilities. This observation correlates with data from a large study conducted by Fichten, Jorgensen, Havel, and Barile (2006) in which factors reported by recent graduates ($N = 1486$) from colleges in Montreal, Canada that served as facilitators or barriers to their success were similar amongst the participants with and without disabilities. Likewise, Lang and colleagues (1993) found that hearing impaired students in their study highlighted similar teacher characteristics as being effective compared to that reported from other studies involving students without hearing impairments.

2.3.2 Potential of e-learning to enhance accessibility. E-learning has the potential to be more accessible than traditional FTF learning. Increased accessibility is possible due to the potential for built-in flexibility and adaptability with respect to when, where, and how learning takes place. These are factors that are often more rigid in traditional FTF course delivery, and which can contribute to barriers to learning. This section of this literature review will explore the ways in which learning technologies may help to reduce or contribute to barriers to learning that post-secondary students may encounter. Because post-secondary students with various disabilities have a wide array of learning needs and preferences, and students without disabilities may also identify with similar needs and preferences in certain contexts, the discussion to follow will draw heavily on the growing collection of literature at the intersection of e-learning, disability, and higher education.

2.3.2.1 Flexibility of when and where e-learning takes place. Enhanced accessibility in e-learning environments may be experienced by all students in certain contexts, including students with various types of visible and non-visible disabilities. For example, students with severe physical disabilities may find the flexibility of when and where learning takes place to be particularly helpful. Students that use wheelchairs may find campuses to be less than fully accessible, even when institutions are in compliance with government-regulated building codes (Ontario Human Rights Commission, n.d.). Working from home or from another accessible environment of their choice is thus an attractive alternative to attending FTF classes on campus. Additionally, students who are uncomfortable sitting upright or sitting still for long periods of time can benefit from studying at home where they can control the length of the intervals that they spend sitting.

Students with non-visible medical or mental health disabilities may also benefit from flexibility of e-learning with respect to when and where learning takes place. Students with mental health disabilities (e.g., depression, bipolar disorder) may experience less anxiety working from home and where they may take additional time to contribute to asynchronous class discussions and email (Grabinger, 2010) versus fast-paced synchronous FTF in-class discussions. Students with medically-related disabilities may also benefit from asynchronous aspects of e-learning. Chronic or episodic illness may necessitate missing scheduled FTF classes, however the opportunity to access course content and contribute to class discussion during periods of wellness is an attractive attribute of e-learning.

Students without disabilities may also appreciate flexibility of when and where e-learning takes place. For example, this flexibility can help to alleviate scheduling challenges that some students including students with dependents and working students may face. Students can choose to devote time to coursework that not only best suits their schedule, but also their preferred learning times. For example, a "morning person" may choose to set aside time for a course early in the day. The ability to work from home also reduces the time and expense involved in commuting to campus for FTF classes or moving to be closer to campus.

2.3.2.2 Flexibility of how e-learning takes place. The use of learning technologies can allow for many possibilities with respect to how students receive, interact with, and express understanding of learning materials. This flexibility can allow students to tailor their learning experiences to their unique learning styles and preferences, and may also help to meet specific needs of students related to their ability/disability or learning context.

In e-learning courses, LMSs are frequently used as platforms to create course websites where course materials are housed in electronic format (Samarawickrema & Stacey, 2007; White & Larusson, 2010). Popular LMSs used in higher education include Blackboard (http://www.blackboard.com/), Desire2Learn (http://www.desire2learn.com/), and Moodle (http://moodle.org/). A recent study found that LMSs, email, and electronic versions of course content were among the most accessible educational technologies (Asuncion, Fichten, & Barile, 2007). For example, electronically available text-based course materials that are compatible with screen reading software (including Word, PowerPoint, or PDF files; text descriptions of images; and transcripts of audio and video files) increase accessibility of content for students with visual and hearing impairments as well as students with learning disabilities. The availability of course content in multiple formats may be helpful for many students, in addition to those students with sensory and learning disabilities. For example, students with various types of disabilities have indicated that the availability of online course notes is helpful (Fichten, et al., 2009).

In addition to facilitating provision of course content in multiple formats, LMSs can also be useful for posting of supplementary materials and for facilitating different modes of communication. Posting of supplementary materials such as outlines, organizers, and note-taking aids can be particularly helpful for students with learning disabilities who may have difficulty with organization and with identifying key points from lectures (Cook & Gladhart, 2002; Gladhart, 2010; Strangman, Hall, & Meyer, 2003). The use of asynchronous communication tools (e.g., email or discussion forums) within an LMS may also be valued by students with learning disabilities who may feel less comfortable participating in FTF classroom discussions and who may benefit from more time to organize their thoughts (Cook & Gladhart, 2002; Palloff & Pratt, 1999), and for *Deaf* students who can participate in class discussions without the need for interpreters. Moreover, students with learning disabilities or visual-motor impairments may find it easier to communicate their understanding of course content on online tests in which there is not a need to alternate between a separate question sheet and answer booklet (Cook & Gladhart, 2002).

2.3.2.3 Summary of e-learning benefits. In summary, the ability to choose when, where, and how to access and interact with course content can increase the accessibility of higher education for all students. FTF learning environments that are less flexible may result in mismatches between the needs of students and the learning environment, while e-learning environments have the potential to reduce these mismatches by supporting the inclusion of more flexibility and thus enhanced accessibility.

2.3.3 Potential of e-learning to create barriers to accessibility. While there is potential for e-learning to remove access barriers and to increase accessibility of higher education, this is not always observed. Adopting the use of new learning technologies allows universities to stay current and to offer courses with a modern aesthetic (Harper & DeWaters, 2008), but accessibility is frequently overlooked when courses are designed (Seale & Cooper, 2010). Incorporation of technology into e-learning courses is problematic when the technology reproduces (or enhances) barriers traditionally encountered in the FTF classroom and/or creates new barriers to access. This can occur as a result of inaccessible features of any aspect of the e-learning infrastructure – a term that Coombs (2010) has used to describe the collection of websites and other technologies and electronic materials that students must access to participate in e-learning.

2.3.3.1 Course LMSs. Both attributes of LMSs themselves and the ways in which they are used by educational institutions and individual teachers may affect the accessibility of the learning environment. Though vendors of LMSs strive to meet accessibility standards and to prepare learning platforms with built-in accessibility features, such features are not helpful if institutions do not make them available for use, and/or if teachers do not know that they are available or how to activate them (Elias, 2010; Gladhart, 2010). For example, in an examination of an online graduate course offered in Moodle, Elias (2010) found that only approximately one third of the features of the LMS that can increase accessibility were utilized. Even with an accessible LMS, the need to learn how to use a sophisticated course management system can also be intimidating for students with disabilities if training and support is not provided (Seale, Draffan, & Wald, 2004).

The availability and use of LMS accessibility features are not the only considerations that affect accessibility. The manner in which content is posted and organized in an LMS can have a large impact on accessibility (Cook & Gladhart, 2002; Gladhart, 2010). If course content is posted by the teacher in a disorganized or cluttered manner, this can be confusing for students including those with learning disabilities (Gladhart, 2010). Seale, Draffan, and Wald (2004) also found that the default layout of

Blackboard was confusing for students with dyslexia who often reported difficulty with finding essential course materials.

A third problem that can arise while using an LMS is when online testing features are used. One reason for this is that extended time on tests is a very common accommodation for students with several types of disabilities (Lovett, 2010), yet it is often not possible to modify online test settings in order to allow selected students additional time (Cook & Gladhart, 2002). Secondly, even if it is possible to extend the test time, time-limited testing in which the timer cannot be stopped to allow students to pause and resume the test is problematic for students who cannot complete a test in one sitting. For example, students who may experience a medical episode such as a seizure during the test will not benefit from additional test time unless the test can be paused and resumed. Thirdly, some students who use assistive software with LMS pages have found that the software is not compatible with certain types of online test questions, and that they may devote more effort to the mechanics of answering a question than to considering their actual response (AFB, 2008; Rangin, 2009).

Additional access issues that can arise as a result of the nature of the course materials posted in an LMS and forms of electronic communication that are supported by LMSs are discussed in the following sections.

2.3.3.2 Text-based content. Students with print disabilities have difficulty accessing text in the same way that non-disabled users do. For example, students with low vision or dyslexia may find certain fonts hard to read and may have difficulty reading text when there is inadequate contrast between text and background colors (DRC, 2004;

Rainger, 2003; Sapp, 2009), and students with dyslexia often find large segments of text overwhelming (Coombs, 2010). Students with motor impairments may also be print-disabled if they experience difficulty physically navigating through computer-based text. For print-disabled students, screen reading software that reads text aloud can overcome challenges associated with accessibility of text-based content (Coombs, 2010; Rainger, 2003). However, webpages and documents are frequently not formatted in a manner that is compatible with screen readers (Coombs, 2010; Fichten, et al., 2009).

Webpages in html format are generally compatible with screen readers but since many webpages are designed with frames (where content is organized in columns), the output from screen readers that read content from left to right can be confusing (Cook & Gladhart, 2002). Similarly, instructor-prepared documents linked within course websites are also frequently inaccessible if they include content organized in columns or in text boxes. In addition to text boxes, PowerPoint presentations often include tables, buttons, and hyperlinks that are ignored by screen readers (Coombs, 2010; Grace & Gravestock, 2009). PDF documents are particularly problematic, as they must be created using the *Style* feature in a word processing program so that the PDF version can be tagged in order to be screen reader compatible – these steps are often omitted by faculty when preparing PDF documents (Gladhart, 2010; Grace & Gravestock, 2009).

2.3.3.3 Electronic communication. E-learning courses often include electronic communication tools that facilitate asynchronous communication (e.g., text-based email or discussion forums) or synchronous communication (e.g., live text-, video-, or multimedia-based chat). In addition to challenges associated with print disabilities described in the preceding section, text-based electronic communication may be difficult for students with learning disabilities or mental health disabilities to interpret and participate in, as some students have impairments that make it difficult to discern emotion from text (Grabinger, 2010; Grace & Gravestock, 2009). This can lead to misunderstandings and even feelings of anger or defensiveness by misinterpreting written remarks from the instructor or classmates as criticisms (Grabinger, 2010).

Text-based communication that takes place in synchronous chat sessions can be particularly hard to follow due to the fast pace, and students with cognitive or learning disabilities may find it difficult to follow a chat session where there are multiple conversations taking place simultaneously (Cook & Gladhart, 2002). Woodfine, Nunes, and Wright (2008) reported that students with dyslexia felt embarrassment and shame at not being able to write high quality text in a synchronous forum, and pressured to try to read faster than they are comfortable with. Students with learning disabilities such as dyslexia or other cognitive disability that affects short-term memory recall may also find it difficult to keep up with synchronous text-based communication (Rainger, 2003; Woodfine, et al., 2008). Additionally, many discussion forums and chat rooms are incompatible with screen reading software, and therefore students with visual impairments may also have difficulty participating in text-based electronic communication (Seale, 2006a).

Live multimedia chat sessions are especially problematic, as there can be access issues with respect to text, audio, video, and interactive elements that may be experienced by students with various types of disabilities. Moreover, the simultaneous use of various modes of communication can be over-stimulating and overwhelming for students with mental health disabilities (Grabinger, 2010). Access issues relevant to multimedia content are discussed further in the following section.

2.3.3.4 Multimedia and Web 2.0 content. Modern e-learning courses look different than traditional distance education courses, as there has been a trend towards replacing text-based content with various forms of multimedia (Armstrong, 2009; Burgstahler, Corrigan, & McCarter, 2004) and growing interest in the use of Web 2.0 tools in education. Multimedia applications include combinations of media such as text, images, video clips, animations, sound, and interactive elements. In addition to multimedia applications that are freely available on the Internet, software companies such as Microsoft offer many technologies that are interactive (e.g., with buttons, or "drag-and-drop" features) and visual (e.g., with photos, graphs, or charts), and which are promoted for use as e-learning materials (Armstrong, 2009). Web 2.0 tools solicit interaction from users, and include blogs, wikis, Facebook, YouTube, and Twitter which are frequently used in higher education (for examples, refer to Huang & Nakazawa, 2010; Kim, Hong, Bonk, & Lim, 2011; Lowe & Laffey, 2011). While multimedia applications and Web 2.0 tools can aid in the development of courses that appear to be sophisticated and of a high quality, these technologies can create significant access barriers for some students.

Incorporation of Web 2.0 tools into e-learning environments may allow for increased flexibility in terms of how information is presented and how students participate. This may be beneficial for students with disabilities and indeed, Grabinger (2010) has suggested that such flexibility may be useful for some students with cognitive disabilities. However, students may encounter incompatibility between Web 2.0 tools and screen reading software. For example, when a Facebook page is updated when a friend posts new content, a screen reader may begin to re-read the entire page from the beginning (Ribera et al., 2009). YouTube has just recently added a captioning function that may be used when new videos are uploaded, but many YouTube videos remain uncaptioned and thus may be inaccessible to students with hearing loss.

The benefits of e-learning for students with sensory disabilities may be diminished if multimedia applications are used in a manner that recreates or intensifies barriers experienced in the traditional FTF classroom. For example, students with visual disabilities are often disadvantaged because many e-learning technologies are vision-centric (Armstrong, 2009). While text-based descriptions can be provided for visual content, use of multimedia applications may create sophisticated graphical representations of course content that cannot be effectively translated to a text-based description (Armstrong, 2009). Students with hearing loss may find it possible to speechread in a FTF class, though the use of multimedia applications with audio and visual components that do not facilitate speechreading and which lack captions or transcripts renders the content inaccessible (Fichten, et al., 2009).

The interactive nature of many multimedia applications that require hand-eye coordination is problematic for students with visual-motor impairments. For example, mouse-driven exercises are not accessible unless they are also compatible with keyboard or voice input (Grace & Gravestock, 2009; Sapp, 2009). If there is a need to use specialized devices such as a head pointer or a foot operated mouse, students will be at a disadvantage unless they are given extra time to complete tasks (Grace & Gravestock, 2009).

2.3.3.5 Websites external to the course LMS pages. The e-learning infrastructure often extends beyond the course LMS pages and the content housed therein. Students must navigate through other institutional websites to reach the LMS pages, and also frequently need to access other academic websites (e.g., the library homepage) and websites external to the institution (e.g., databases) to complete their studies. Unfortunately, there is widespread inaccessibility across the Internet including academic and non-academic websites (DRC, 2004; Harper & DeWaters, 2008; Thompson, Burgstahler, & Moore, 2010). A myriad of Web accessibility issues may be encountered, with frequently reported issues including confusing page layout and/or navigation, absent or uninformative ALT tags (alternative, ALT, text describing what is depicted) for images, poor text-to-background contrast, small text and graphics, and incompatibility with assistive technologies.

2.3.3.6 Summary of e-learning barriers. In summary, e-learning infrastructure is frequently inaccessible to students, particularly students with disabilities. Problems range from websites that are confusing and difficult to navigate, to unfair testing procedures, to inaccessible means of communication, and to content that is presented in inaccessible forms. Even with the use of assistive technologies such as screen readers and built-in accessibility features of LMSs, it is often not possible for students to access all e-learning infrastructure and course content or to participate in all forms of electronic communication. There are potential accessibility problems associated with all components of typical e-learning infrastructure.

2.4 Designing E-Learning Environments with Accessibility in Mind

Following a review of the benefits and drawbacks of e-learning for students with disabilities, it is clear that e-learning has the potential to enhance accessibility of higher education for some students yet often does not live up to this potential. Accessibility is a complicated issue because attributes of e-learning infrastructure that enhance accessibility for one student may decrease accessibility for another. For example, text-based content may be very accessible for a *Deaf* student who can access the content without an interpreter, yet the same content may be less accessible than auditory content for a dyslexic student or a non-disabled student who understands better from listening than from reading.

How then, can accessible e-learning environments be designed? It is useful to draw upon a definition of accessible learning technologies that supports the preceding discussion. Consider these remarks from the IMS Global Learning Consortium, in which accessibility and disability are defined in the context of e-learning:

> The term disability has been re-defined as a mismatch between the needs of the learner and the education offered. It is therefore not a personal trait but an artifact of the relationship between the learner and the learning environment or education delivery. Accessibility, given this re-definition, is the ability of the learning environment to adjust to the needs of all learners.... The needs and preferences of a user may arise from the context or environment the user is in, the tools

> available..., their background, or a disability in the traditional [medical model] sense. Accessible systems adjust the user interface of the learning environment, locate needed resources and adjust the properties of the resources to match the needs and preferences of the user. (IMS Global Learning Consortium, 2004)

This description of accessibility emphasizes the need for flexible technology and suggests that the burden of adaptability should shift from the student to the e-learning environment. Therefore, if e-learning infrastructure is to be accessible to a student population with a wide range of needs, it must be sufficiently flexible.

2.4.1 Creating accessible electronic course materials. Course materials may be designed with flexibility in mind so that they may be utilized in different ways. Text-based course content must be compatible with screen reading software, and text equivalents must be provided for visual and audio content. Methods for creating accessible PDF and Microsoft Office documents and presentations are outlined in several books and online resources (for example, refer to Coombs, 2010; Grace & Gravestock, 2009; Seale, 2006a; WebAIM, n.d.). For example, the *Styles* feature of Microsoft Word must be used in order to assign headings (and thus structure) to a document. This feature allows for what computer programmers refer to as "semantic markup" that can be used to define attributes of electronic content (e.g., header text). Without clues about the structure of a text-based document offered by semantic document markup, text-to-speech output from screen reading software may be confusing. ALT text describing what is depicted in an image must be provided for all images (the steps required to do so vary according to the type and version of program being used) unless they are decorative only, in which case null text (a space or “”) must be included to signal to a screen reader to ignore the images. These steps must also be taken prior to converting Microsoft Word documents into PDF documents. When creating PowerPoint presentations, ensuring that all essential text is visible within the Outline view is recommended, as text boxes and text within other graphics created within PowerPoint is generally ignored by screen readers.

2.4.2 Creating or selecting accessible websites. Whether creating course webpages within an LMS, creating an external website, or selecting websites to direct students to visit, there are basic steps that can be taken to conduct a preliminary evaluation of the accessibility of the pages. According to the World Wide Web Consortium (W3C, 2005) and Utah State University's Web Accessibility in Mind Initiative (WebAIM, 2009), useful preliminary accessibility evaluation steps include: a) ensuring that all images have ALT text and that it is succinct and useful, b) enlarging the font using Web browser controls to ensure that pages remain readable with minimal horizontal scrolling, c) testing keyboard navigation through the page (e.g., using the Tab key rather than a mouse), d) confirming the availability of text equivalents for audio content, and e) ensuring that hyperlink text is informative by indicating what the link is for. While these steps are helpful as a quick preliminary test of accessibility, a more comprehensive examination of websites is required in order to more accurately assess accessibility. Web accessibility guidelines have been developed for this purpose, and are described in the following section.

2.5 Web Accessibility Guidelines

The need to consider how to enhance accessibility of the Web for people with disabilities was recognized shortly after its creation in 1990. Ellcessor (2010) speculates that the development of Web accessibility guidelines in the U.S. was driven by several events taking place at this time, including increased attention being given to the civil rights of people with disabilities with the passage of the Americans with Disabilities Act (ADA) in 1990, development of the first Web browser (Mosaic) in 1993 that supported graphical and multimedia display of content (not compatible with assistive technologies at the time), and increased popularity of the Web. In 1995, the first Web accessibility guidelines were published by the Trace Center of the University of Wisconsin-Madison (Trace Center, 2007), and were intended to assist in development of webpages that were accessible when viewed with Mosaic (Vanderheiden, 1995). A subsequent "unified" version of guidelines developed by the Trace Center that incorporated guidelines developed by several other agencies was ultimately used as a starting point for development of the Web Content Accessibility Guidelines (WCAG) (Trace Center, 2007). The WCAG have emerged as an internationally recognized and influential set of guidelines.

2.5.1 Development of the WCAG 1.0. In 1994, the World Wide Web Consortium (W3C) was founded by Tim Berners-Lee, one of the inventors of the Web, to promote its continued growth by development of relevant protocols and guidelines (W3C, 2009a, 2009b). Three years later, the Web Accessibility Initiative (WAI) was chartered as a W3C initiative aimed at enhancing Web accessibility for people with disabilities

(Dardailler, 2009). In August of 1997, a WAI working group was formed to develop a comprehensive set of Web accessibility guidelines and supporting documentation that would become the WCAG version 1.0 (WAI, 1997). Subgroups were formed to focus on different aspects of Web accessibility (Ellcessor, 2010).

During the process of developing the guidelines, aspects of accessibility, usability, and universal design were considered (Ellcessor, 2010). By including people with disabilities within the guideline working group, user-centered design was also practiced (Ellcessor, 2010). While the WCAG are referred to as accessibility guidelines, the working group chose to link accessibility and usability within the guidelines and associated documentation. For example, while navigability is an attribute typically associated with usability, information and guidelines relevant to navigability also appear in the WCAG (W3C, 1999). This merging of accessibility and usability reflects efforts of the working group to develop guidelines that would be representative of the needs of users with a large range of preferences and needs under a variety of circumstances, a concept which later came to be referred to as universal usability.

Universal usability has been formally defined as "a focus on designing products so that they are usable by the widest range of people operating in the widest range of situations as is commercially practical" (Vanderheiden, 2000, p. 32). This is similar to the concept of universal design which is referred to as "the design of products and environments to be usable by all people, to the greatest extent possible, without the need for adaptation or specialized design" (Center for Universal Design, 1997). Interestingly, Vanderheiden (co-chair of the WCAG working group) not only published the definition and description of universal usability (Vanderheiden, 2000) but also contributed to the development of the principles of universal design that were originally developed for architecture (Center for Universal Design, 1997).

The WCAG 1.0 were published on May 5, 1999 along with supporting documentation including descriptions of Web accessibility access issues and challenges that people with different types of disabilities may face, themes of accessible design, and an overview of how the guidelines are organized and prioritized and how they may be applied to conformance testing (W3C, 1999). The WCAG 1.0 consists of 14 guidelines which include a description of the rationale and benefits, along with a list of relevant checkpoints and examples of techniques that may be used to meet each checkpoint.

The checkpoints associated with each guideline were assigned priority levels by the working group to indicate the relative importance. Priority 1 checkpoints must be satisfied or a webpage will be inaccessible to some users; Priority 2 checkpoints should be satisfied or some users will have difficulty accessing content on the webpage; and satisfying Priority 3 checkpoints will make it easier for some users to access content. For example, consider Guideline 2 – Don't rely on color alone. The corresponding checkpoint 2.1, namely, ensure that all information conveyed with color is also available without color (for example from context or markup), is deemed Priority 1, and one suggested technique for meeting this checkpoint is to use header elements to convey page structure. Checkpoint 2.2, namely, ensure that foreground and background color combinations provide sufficient contrast when viewed by someone having color deficits or when viewed on a black and white screen, is deemed Priority 2 for images and Priority 3 for text. If a webpage is examined for compliance with checkpoints for all guidelines, corresponding levels of conformance may be assigned: Level A sites meet all Priority 1 checkpoints, Level AA sites meet all Priority 1 and 2 checkpoints, and Level AAA sites meet all Priority 1, 2, and 3 checkpoints.

The WCAG 1.0 have been adopted as international benchmarks for Web accessibility and have influenced legislation in several countries including Canada (refer to the *International legislative influence of the WCAG* section below). However, the availability of new technologies that are not addressed by the WCAG 1.0, coupled with the subjective nature of several guidelines that render conformance testing and enforcement difficult, made it necessary to work towards a new version of the guidelines shortly after version 1.0 was released (Reid & Snow-Weaver, 2008). Indeed, the WCAG 2.0 working group was formed in 2000 and the first public working draft of the WCAG 2.0 was published on January 25, 2001 (W3C, 2001).

2.5.2 The WCAG 2.0. Primary goals of the WCAG 2.0 working group were to revise the version 1.0 guidelines such that they were reliably testable in recognition of their likely use as standards for accessibility conformance testing, and "technology-neutral" to reflect the current diversity of Web content and its ever-evolving nature (Reid & Snow-Weaver, 2008). Following the release of the first working draft of the WCAG 2.0 in 2001, nearly eight years passed before consensus was reached and a stable version endorsed by the W3C was released on December 11, 2008 (W3C, 2008).

The WCAG 2.0 consists of four principles, namely, that Web content should be a) perceivable: "information and user interface components must be presentable to users in ways they can perceive;" b) operable: "user interface components and navigation must be operable;" c) understandable: "information and the operation of user interface must be understandable;" and d) robust: "content must be robust enough that it can be interpreted reliably by a wide variety of user agents, including assistive technologies" (W3C, 2008, WCAG 2.0 Guidelines section). These principles are very similar to Vanderheiden's description of the basic components of universal usability (Vanderheiden, 2000, p. 36) and, as such, the overlap between accessibility and usability is retained in the version 2.0 guidelines. This linkage is further strengthened by the inclusion of additional checkpoints related to navigation and user control of the Web browsing experience (Ribera, et al., 2009).

Specific success criteria classified as Level A, AA, or AAA are associated with each guideline, and there are 61 criteria in total. The success criteria have been written as testable statements such that they may be used to evaluate guideline conformance. The W3C recommends that conformance schedules replace the WCAG 1.0 with the WCAG 2.0 (W3C, 2008). In contrast to the WCAG 1.0 where criteria were assigned priority levels, the WCAG 2.0 levels imply suggested implementation order rather than relative importance (Reid & Snow-Weaver, 2008).

Websites may be deemed fully Level A, Level AA, or Level AAA WCAG 2.0 conformant, in a similar manner to the method used previously with the WCAG 1.0: Websites that meet all of the Level A success criteria are Level A conformant; websites that meet all of the Level A and AA success criteria are Level AA conformant; and Level AAA conformant websites meet all of the WCAG 2.0 A, AA, and AAA success criteria.

Websites may also be deemed partially conformant. Partial conformance may apply if the webmaster has met all of the success criteria at a particular level but has not created all of the content on the webpage. For example, if a webmaster has created a fully Level AA conformant webpage that includes integrated Web 2.0 tools, users may contribute content that appears on the website which is not accessible (Reid & Snow-Weaver, 2008) and the website would become partially Level AA conformant.

In order to remain technology neutral, the stable main WCAG 2.0 document lacks technology-specific language and does not prescribe specific steps that must be taken to meet the guidelines (Reid & Snow-Weaver, 2008). To remain current with emerging technologies, a supplementary *Techniques* document that is expected to be updated frequently with current technology-specific suggestions is available (W3C, 2008). A second supplementary *Understanding* document is available to assist with understanding and implementation of the guidelines.

2.5.2.1 Limitations of the WCAG 2.0. Since their release in 2008 (and prior to this, based on working drafts), criticism of the guidelines has emerged from within and external to the WCAG 2.0 working group. Primary criticisms are the lengthy documentation, ambiguous language, and obscure jargon (Clark, 2006, May 23; Kapsi, Vlachogiannis, Darzentas, & Spyrou, 2009a, 2009b; Ribera, et al., 2009). Insight from members of the working group suggests that the ambiguous text of the root document supports its technology-neutral stance, the supplemental *Understanding* document is more educational than WCAG 1.0 documentation, and the supplemental *Techniques* document will keep the WCAG 2.0 current via frequently-updated technology-specific suggestions

(Reid & Snow-Weaver, 2008; Termens et al., 2009). Collectively, however, these three documents yield in excess of 400 printed letter-sized pages (Clark, 2006, May 23) and are frequently described as difficult to understand (Alonso, et al., 2010; Kapsi, et al., 2009b; Seeman, 2006b). Clark, a member of the working group and an outspoken critic of the guidelines, expressed concern that the *Understanding* document is more than twice as lengthy as the root document (Clark, 2006, May 23). Seeman, also a member of the working group, has indicated that the *Techniques* document is not accessible to her even with extensive knowledge of the guidelines, due to limitations that she experiences associated with a learning disability (Seeman, 2006b). While hyperlinks direct readers to definitions of complex terminology, the definitions themselves may not be accessible to some users (refer to Clark, 2006, May 23, for more examples).

Additional concerns are related to testability of the success criteria, which was one of the key intended improvements of the WCAG 2.0 compared to version 1.0. According to the W3C, "WCAG 2.0 success criteria are written as testable statements" and reliably human testable is defined as testable "by human inspection and it is believed that at least 80% of knowledgeable human evaluators would agree on the conclusion" (W3C, 2008). Concern has been raised about the potential for success criteria checklists to support conformance testing that may not effectively pinpoint accessibility or usability problems (Brajnik, Yesilada, & Harper, 2012; Kapsi, et al., 2009b; Kelly, Sloan, Phipps, Petrie, & Hamilton, 2005; Ribera, et al., 2009). For example, the presence of ALT text for an image may be satisfied even if the text is not informative to users. Moreover, a few small-scale studies have been published in which results of WCAG 2.0 conformance testing by different evaluators are compared. Such studies have found that many of the success criteria are not reliably human testable (Alonso, et al., 2010; Brajnik, 2009; Brajnik, et al., 2012) and that some criteria that are generally deemed ambiguous were actually found to be more reliable than those generally deemed more straightforward and vice versa (Alonso, et al., 2010; Brajnik, 2009; Brajnik, et al., 2012). Given the implications of this preliminary data for the appropriateness of basing conformance testing on the WCAG 2.0, further investigation into criteria testability is essential.

A related issue is the tiered structure of the success criteria. While the 2.0 level designations are intended to represent suggested implementation order, it is not clear whether that intent will translate to legislative use and mandated website conformance. For example, in 2012 the Federal Government of Canada updated Web accessibility policy relevant to all webpages provided through Government of Canada websites. Current policy mandates gradual phasing in of WCAG 2.0 AA conformance, omitting mandatory compliance to Level AAA criteria (Treasury Board of Canada Secretariat, 2011). Similarly, the Integrated Accessibility Standards of the AODA in the province of Ontario mandates conformance of newly built or substantially modified websites to Level A and AA WCAG 2.0 criteria, and AA compliance will not be required until 2021 (AODA Integrated Accessibility Standards, 2011). Given that Level AAA criteria may be very important to users with some disabilities, omission of AAA criteria in compliance schedules is troublesome.

2.5.2.1.1 The WCAG 2.0 and learning disabilities. Web accessibility is often particularly low for individuals with learning disabilities compared to other disability groups, as a result of the complexity associated with this category of disability (Friedman & Bryen, 2007; Laff & Rissenberg, 2007; McCarthy & Swierenga, 2010; Nicolle & Paulson, 2004). While the WCAG 2.0 indicate that they address, at least in part, the needs of users with learning disabilities, this is accompanied by the acknowledgement that full conformance may not necessarily ensure accessibility (W3C, 2008). Even this tentative suggestion of the value of the WCAG 2.0 for these disability groups has raised objection (Seeman, 2006a) and the minimal representation of related experts within the working group has also been highlighted (Kennedy, et al., 2011). Recall that a member of the working group who identified herself as having a learning disability found the *Techniques* document inaccessible even though it was prepared according to the WCAG 2.0 (Seeman, 2006b).

While Web accessibility needs for users with learning disabilities may not be easily understood, there are common Web-related difficulties that are experienced by users who identify as persons with learning or cognitive disabilities. A large-scale study examining accessibility of websites that included dyslexic user participants found frequently encountered difficulty associated with confusing page layout and navigation, inappropriate use of colors, poor content to background contrast, small graphics and text, and complicated language (DRC, 2004). When these findings are combined with WCAG 2.0 working group documentation related to meeting the needs of users with cognitive and learning disabilities (Vanderheiden, 2006), the finalized guidelines themselves, and criteria testability studies (Alonso, et al., 2010; Brajnik, 2009), the following findings emerge:

1. Two-thirds or more of the WCAG 2.0 success criteria may be helpful for users with learning disabilities;
2. Common problems experienced by dyslexic Web users relate to criteria from all three levels (minimizing the usefulness of legislation mandating only Level A and AA conformance); and
3. Several Level A criteria that may be useful towards addressing many needs of those with learning disabilities may not be reliably testable.

2.5.2.2 Summary. In summary, while the WCAG 2.0 are helpful in enhancing Web accessibility, accessibility and usability are not necessarily concomitant with website conformance. A significant time commitment to reading and understanding the documentation is required of designers and conformance testers alike. Moreover, input from potential users is an essential additional step towards identifying whether a conformant website is truly accessible and usable for the intended audience.

2.5.3 International legislative influence of the WCAG. Several countries have either incorporated the WCAG themselves into Web accessibility policies or legislation, or have created national guidelines or mandatory standards that have been inspired by the WCAG. For example, the Australian government mandated that all government websites conform to WCAG 2.0 Level A by December 2012, and Level AA by December 2014 (Australian Government, 2011). In Canada, three phases of implementation of WCAG 2.0 conformance included mandatory Level AA conformance for all Federal government websites by July 2013 (Treasury Board of Canada Secretariat, 2011). Ontario is the first province in Canada to adopt provincial accessibility standards. The Integrated

Accessibility Standards of the Accessibility for Ontarians with Disabilities Act (AODA) include a conformance schedule for websites of the provincial government, public sector, and large organization websites. This standard mandated gradual phasing in of WCAG 2.0 conformance up to Level AA by January 2012 for new government websites and January 2021 for public sector and large organization websites (AODA Integrated Accessibility Standards, 2011).

There are many examples of countries that have adopted variations of the WCAG. This includes Germany, Hong Kong, Italy, Japan, and Sweden (Thatcher et al., 2006, pp. 546-579). The U.S. is another example of a country that has developed its own WCAG-inspired Web accessibility standards, namely, the Section 508 standards of the Rehabilitation Act. The Section 508 standards deserve special attention, as the U.S. was the first country to adopt mandatory accessibility standards and these standards are also widely recognized and adopted internationally.

2.5.3.1 Section 508 Standards of the U. S. Rehabilitation Act. The U.S. Rehabilitation Act is a piece of legislation intended to prevent discrimination against people with disabilities by federal agencies (Rehabilitation Act, 1998). Section 508 was added to the Rehabilitation Act in 1986 and, following subsequent amendment, Web accessibility standards went into effect in 2001. As such, the Section 508 standards are applicable to websites of Federally-funded agencies.

The Section 508 standards include 16 rules that apply to Web-based intranet and Internet information and applications (Rehabilitation Act, 2000). Of these rules, the first 11 correspond to WCAG 1.0 Priority 1 checkpoints, while the remaining 5 rules are unique to Section 508. Wakefield, an author of the standards, has described the rationale behind using the WCAG 1.0 as a starting point but not simply adopting them (Ellcessor, 2010; Wakefield, 2000). Development of the standards began in 1998, which was the same time that the WAI was finalizing the WCAG 1.0. According to Wakefield, there was pressure on the Section 508 standards working group to adopt the WCAG 1.0 because the W3C's WAI (refer to section *2.5.1 Development of the WCAG 1.0* section for a discussion of the W3C and the WAI) that was charged with developing the WCAG was Federally-endorsed (Ellcessor, 2010). However, concerns about ambiguity/subjectivity of some of the WCAG 1.0 checkpoints that may render enforcement difficult, and different opinions regarding accessibility, usability, and universal design (unlike the WCAG working group, the Section 508 working group preferred to separate accessibility from usability and expressed skepticism about the value of universal design), led to the removal of some of the WCAG 1.0 guidelines (Ellcessor, 2010; Wakefield, 2000). The Section 508 standards working group did, however, consult with the WAI who reviewed and offered feedback during the process of developing the Section 508 standards (Ellcessor, 2010).

2.5.4 Relevance of Web accessibility guidelines to e-learning in higher education. In countries or jurisdictions where there is accessibility legislation that is applicable to higher education, conformance of Web content to the WCAGs (or standards inspired by them) may be legally enforceable. One such example is the U.S. where the ADA and the Rehabilitation Act make it unlawful for agencies that are publicly or government funded, respectively, to provide electronic or information technology that is inaccessible to people with disabilities (Americans with Disabilities Act, 1990; Rehabilitation Act, 1998). Depending on their source(s) of funding, post-secondary institutions in the U.S. are therefore subject to one or both of these Acts. The U.S. Department of Justice ruled in 1996 that the ADA applies to the Internet (Thatcher, et al., 2006, p. 515), and recommends that Web developers refer to the Section 508 standards and the WCAG for guidance while creating websites for agencies funded by state or local governments (U.S. Department of Justice, 2008). In addition, Section 504 of the Rehabilitation Act stipulates that persons with disabilities may not be prevented from participating in any Federally-funded program or activity. The U.S. Department of Education, Office of Civil Rights has explicitly indicated that Section 504 applies to all e-learning infrastructure including online courses, any online content, and emerging technologies, including those selected for pilot testing (Joint Department of Justice and Department of Education, 2011).

In Canada there are no Federal requirements for accessibility of non-government content, and Web accessibility is under provincial jurisdiction. In the province of Ontario, the Integrated Accessibility Standards of the AODA that mandate WCAG 2.0 Level AA conformance are applicable to public sector and large organizations (AODA Integrated Accessibility Standards, 2011), and thus to provincial universities and colleges. It will be necessary for post-secondary institutions in the province to ensure that new webpages (and content posted therein, including Web-based applications) are WCAG 2.0 Level A compliant by 2014 and Level AA compliant by 2021.

Vendors of learning technologies are responding to the influence of the WCAG on accessibility policy and legislation. For example, vendors of popular LMSs refer to the

WCAG in accessibility policies. Blackboard (http://www.blackboard.com) follows Section 508 and the WCAG, and expresses a commitment for continued user testing and accessibility improvement; Moodle (http://moodle.org/) plans conformance with Section 508, WCAG, and the U.K. SENDA legislation (Special Educational Needs and Disability Act 2001, 2001, c. 10) and invites the public to contribute to open online discussion forums regarding continued improvements related to Moodle accessibility; and Desire2Learn (http://www.desire2learn.com/) follows the WCAG and other WAI guidelines, with conformance checklists publicly available on the website. It is likely that vendors of all technologies that are designed with educational applications in mind will seek WCAG and/or other accessibility guideline conformance in order to maintain relationships with post-secondary institutes.

Legal precedence demonstrating the need for accessible e-learning technologies has been set in the U.S. where the National Federation of the Blind and American Council for the Blind sued Princeton and Arizona State universities for violation of the ADA and Rehabilitation Act for launching pilot programs using the Kindle DX, an e-book reader (U.S. Department of Justice, 2010a, 2010b). Settlement agreements in both cases acknowledge that the e-reader is not accessible for students with visual impairments, and stipulate that the universities may not purchase any electronic devices to be used in classroom settings unless it is accessible to students with visual impairments or an equivalent accessible technology such as an alternate form of e-reader is provided.

2.6 E-Learning Accessibility Testing

There is a scarcity of peer-reviewed studies in the literature describing accessibility testing of e-learning infrastructure. While the aforementioned legal action against Princeton and Arizona State Universities may serve as examples of empirical testing of e-learning accessibility, few relevant peer-reviewed studies could be located. One such study was conducted by Power and colleagues (Power, et al., 2010), in which accessibility testing of LMSs based on WCAG 1.0 compliance was carried out. This research was an empirical study on accessibility of the Moodle, LRN eLearning Platform, and Blackboard LMSs. In this study, expert evaluators conducted "accessibility audits" by examining a variety of pages of each LMS for compliance with ten WCAG 1.0 checkpoints. The expert evaluations were then followed by user testing, wherein four blind students were asked to complete a number of tasks using each platform while using screen reading assistive technology. Both the expert and user evaluations revealed accessibility problems with each of the LMSs tested, and none of the LMSs exhibited Level A WCAG 1.0 compliance. It is surprising that there is not a larger body of research published on this topic, however it is possible that accessibility studies have been conducted by vendors of learning technologies that have not been publicly released. Given the recent finalization of the WCAG 2.0 (in December 2008) and heightened attention given to legal requirements for accessibility of learning technologies (particularly in the U.S. where a number of discrimination lawsuits have arisen and the Federal Government has been explicit about the need for post-secondary institutions to comply with anti-discrimination disability laws), it is likely that this is an area of research that will assume a higher profile in the literature in the near future.

Studies related to student use of e-learning infrastructure that are currently prevalent in the literature are those related to library usability. As academic libraries have begun to offer an increasing number of digital resources, in part fuelled by increased prevalence of e-learning and off-campus library use (Thomsett-Scott, 2004), interest in usability has also increased (Comeaux & Schmetzke, 2007). There are many peer-reviewed studies in the literature that focus on usability testing of library homepages, catalogs, and databases with experimental designs that have included soliciting input from student users (for examples, refer to Bury & Oud, 2004; Cobus, Dent, & Ondrusek, 2005; Denton & Coysh, 2011; Dermody & Majekodunmi, 2010; Jung, Herlocker, Webster, Mellinger, & Frumkin, 2008; Persson, Langh, & Nilsson, 2010). Given the availability of literature on e-learning infrastructure usability testing, the overlap between accessibility and usability within the WCAG, and the fact that usability testing is a well-established field, it is useful to further explore the relationship between accessibility and usability and to examine usability testing methods that may be applied to e-learning accessibility testing.

2.6.1 Relationship between accessibility and usability. The Web accessibility literature often does not clearly distinguish between accessibility and usability, and differing views have been expressed as to the manner in which these attributes are related to one another (Kapsi, et al., 2009b). Arguments have been offered in support of a) accessibility being considered a component of usability, b) accessibility and usability being considered separate attributes, and c) accessibility and usability as separate but overlapping attributes.

Frequently-cited definitions of usability are those offered by the ISO (International Organization for Standardization – an international network of standards institutes) and Jakob Nielsen (a world renowned usability expert). The ISO standard 9241, developed by human-computer interaction experts, defines usability as "The extent to which a product can be used by specific users to achieve specified goals with effectiveness, efficiency, and satisfaction in a specified context of use" (ISO 9241, 1998). Nielsen (1993) suggests that usability typically includes the following attributes: learnability, efficiency, memorability, low error rate, and user satisfaction (p. 26). Offering a simpler approach to describing usability, Rubin and Chisnell (2008) suggest that something is usable when a user does not experience frustration when attempting to use it, and that "the user can do what he or she wants to do the way he or she expects to be able to do it, without hindrance, hesitation, or questions" (p. 4).

Henry (2007) has argued that accessibility is a special case of usability. This viewpoint positions accessibility as an attribute that is nested within the larger notion of usability. Drawing on the ISO 9241 definition of usability, she suggests that the term accessibility applies when the users are people with disabilities and the context of use includes use with assistive technologies (Henry, 2007). In other words, accessibility arises when usability is effective, efficient, and satisfying for more users (including those who identify as persons with disabilities).

Criticisms of this viewpoint are based on observations that accessible systems do not necessarily meet the criteria for usability, and vice versa. Recall that the Section 508 standards working group chose to depart slightly from the WCAG 1.0 in an effort to separate accessibility from usability because it was felt that usability was a more technical attribute that may be assessed in a different manner than accessibility (Ellcessor, 2010). The RNIB (Royal National Institute of Blind People, a charitable organization in the UK supporting people with visual impairment) also supports a distinction between accessibility and usability for this reason (RNIB, 2008). The RNIB offers a helpful example to explain this stance: A webpage may have 40 buttons that are keyboard accessible because they can be accessed using the Tab button. As such, the page is accessible to users that must use a keyboard rather than a mouse to access content. However, the page is not usable because it would be inefficient for a user accessing the page with a keyboard if the most frequently-used button requires the user to press the Tab key 25 times in order to reach it (RNIB, 2008). On the other hand, a webpage may meet the criteria for usability but would be inaccessible to some users if it was not compatible with assistive technologies.

A third perspective is to consider accessibility and usability to have both unique and overlapping attributes. For example, Kapsi and colleagues (Kapsi, et al., 2009b) propose the idea of "usable accessibility" as the intersection between usability and accessibility. In this model, usability includes learnability and memorability; accessibility includes perceivability, operability, and understandability; and usable accessibility, which includes effectiveness and efficiency, represents the "grey area" in the center where usability and accessibility intersect (Kapsi, et al., 2009b). Applying this model, a system exhibits usable accessibility (i.e., effectiveness and efficiency for the user) if users can easily learn and remember how to use it (i.e., it is usable) and users can perceive, operate, and understand the system (i.e., it is accessible). This notion of usable accessibility is similar to Vanderheiden's description of universal usability discussed previously, as a system that allows for use with a wide range of ability and circumstance (Vanderheiden, 2000). Given the blending of accessibility and usability within the WCAG 2.0 and the frequent appearance of the WCAG 2.0 (or related guidelines) within Web accessibility policy and legislation worldwide, this third perspective on accessibility and usability is particularly useful. Moreover, methods employed in the well-established field of usability testing may also prove useful for accessibility testing.

2.6.2 Usability testing. Usability testing is conducted in order to identify problematic areas, critical incidents (incidents that are very positive or very negative), and suggested solutions from users as they work with a product such as an application or a website (Nielsen, 1993). As such, usability testing is a component of a user-centered design process, because input from potential users of the product is sought as it is being developed (Nielsen, 1993; Rubin & Chisnell, 2008). Usability testing may take place at various stages of product development, and testers may interact with a product mock-up (e.g., images, print-outs, or a 3D model of what a product may look like), prototype (e.g., a semi-functional website), or the product in the final stages of development (Rubin & Chisnell, 2008).

2.6.2.1 Types of usability testing. The stage of product development generally influences the method of testing. Exploratory testing (also referred to as formative testing) takes place early on in the product design process, and is carried out in order to get a general idea as to whether or not the proposed design of a product will be usable prior to putting in a substantial development effort (Rubin & Chisnell, 2008). For example, potential users may be asked to view a depiction of a website layout in order to comment on the organization. There is generally a lot of interaction between the user and the test administrator when testing is carried out at this stage. Assessment testing (also referred to as summative testing) takes place further along the development process, and involves the user completing a series of tasks using a product prototype (Rubin & Chisnell, 2008). There is typically less user-administrator interaction during this type of usability testing, as the administrator observes the user as he/she completes the tasks and collects data as the test proceeds. Validation testing (also referred to as verification testing) takes place late in the development phase, and involves the user completing tasks with the product while the test administrator observes in order to identify if previous problems have been addressed (Rubin & Chisnell, 2008). This type of testing may also compare the product usability against previously determined benchmarks (e.g., ensuring that a particular task can be completed in a specified time period). Like assessment testing, there is generally little to no user-administrator interaction during validation testing.

Within each type of usability test, between-subject testing or within-subject testing may be employed. Between-subject testing takes place when different users are asked to test different systems or different components of the same system, while within-subject testing takes place when all users test all systems or all components of the same system (Nielsen, 1993; Rubin & Chisnell, 2008). For example, consider usability testing of a website that aims to examine two distinct sections of the site. Task A requires users to interact with section A of the website, while Task B requires users to interact with section B of the website. If between-subject testing is employed, users will complete either Task A or Task B, while a within-subject test design would require that all users complete both of the tasks. There are advantages to both approaches of assigning tasks to users. Within-subject testing allows for comparison of completion of different tasks by the same user, thus controlling for user variability with respect to speed or prior expertise. However, between-subject testing prevents unwanted influence of learning that could take place if the same tester develops and carries over product expertise from one task to the other. A common way of reducing the impact of this problem when conducting within-subject testing is to use a counterbalanced design, whereby half of the testers complete Task A first and the other half of the testers complete Task B first (Nielsen, 1993; Rubin & Chisnell, 2008).

2.6.2.2 Number of testers required. Nielsen's suggestions about the number of testers required for usability testing are widely cited and generally supported by the literature. He has worked out a pay-off ratio curve that demonstrates that when the number of testers exceeds four, the benefits of testing may no longer outweigh the costs (Nielsen, 1993, p. 174). Along with Landauer, he has also worked out a mathematical formula that can be used to relate the number of users involved in a test to the percentage of usability problems that they are likely to find (Nielsen & Landauer, 1993). To develop the formula, Nielsen and Landauer first examined a collection of usability testing studies that had been previously conducted, and determined that a single user is likely to find 31% of all usability problems. These data were used as the basis of the formula which indicates that five users may find up to 85% of problems and at least 15 users are needed to find all problems (Nielsen, 2000). Nielsen therefore suggests that testing with four to five users is likely to be cost effective and to find the majority of usability problems.

In spite of the popularity of Nielsen and Landauer's model, some experts in the field suggest applying the four to five users suggestion with caution. For example, Rubin and Chisnell (2008) recommend testing with at least eight users. This is because even if four to five users find up to 85% of the usability problems, the remaining unidentified problems may be critical (p. 126). Indeed, Spool and Schroeder (2001) found that the first five users in their usability testing study identified only 35% of the problems within a commercial website, and that the 13th and 15th users each found one unique critical problem that prevented them from completing the task. Also skeptical of the use of only four to five testers, Faulkner (2003) designed an experiment to test the accuracy of Nielsen and Landauer's prediction. She conducted usability testing with 60 users, and then created several data sets including random samples of data from sets of five users in order to determine the percentage of problems that were found. She found that the percentages of problems found by random sets of five users ranged from 55% to 99%. When the data sets were expanded to include random sets of 10 users accuracy increased to a minimum of 80%, and when data from 20 users was included accuracy rose to 95%.

Faulkner (2003) and Woolrych and Cockton (2001) suggest that it is important that researchers understand what Nielsen and Landauer's model means, as well as assumptions that the formula represents when deciding how many users are necessary for usability testing. The four to five user suggestion does not indicate that this number of users will always find the majority of problems, just that this is likely. The estimation of the percentage of problems that a single user will find that Nielsen and Landauer have used when applying their formula was based on 13 studies which, according to Faulkner (2003), was a small sample size and may also invalidate the method that was subsequently used to develop the formula. In other words, Faulkner critiques the statistical methods that were used to arrive at the four to five user suggestion. Woolrych and Cockton (2001) also criticize the statistics as well as assumptions that the formula relies upon. The formula suggests that the probability of finding all problems is equally weighted (i.e., all problems are equally easy or difficult to identify) and that all users would be expected to find a similar number of problems (i.e., that user variation is not a significant factor). Moreover, the model assumes that problems are independent of each other (i.e., a potential "domino effect" of one problem leading to another is not allowed for).

It is logical to expect that characteristics of the users, the tasks, and the application to be tested will influence the likelihood of finding individual problems. As such, while Nielsen and Landauer's suggestion of four to five users is a helpful guideline, care should be taken to consider the effect that attributes of the methodology as well as the desired outcomes of testing (e.g., which will vary according to the type of test, such as formative, summative, or validation) are factored in when deciding how many users to test in a particular context. Nielsen himself has indicated that "Users have infinite potential for making unexpected misinterpretations of interface elements and for performing their job in a different way than you imagine" (Nielsen, 1993, p. 10) and thus it stands to reason that the more users that are included in a usability testing study the more informative the study is likely to be. This may be particularly true when the user population is expected to be diverse in terms of the ways that they interact with technology (e.g., when the intended users of a product include persons with disabilities).

2.6.2.3 Data collected during usability tests. Data collected during usability testing are referred to as performance data if they are related to how a user performs when working with the product under study or preference data if they are related to the user's opinion about the product under study (Rubin & Chisnell, 2008).

Performance measures are typically quantitative, such as the length of time it takes a user to complete an assigned task, the number of tasks completed successfully, the number of error messages a user receives, the number of mouse clicks used to complete a task, etc. This type of data may be collected by the test administrator and/or observers that are present during the test, by viewing video recordings taken while the test took place, or by use of automated usability testing software such as data logging software that runs concurrently during the task performance (refer to section *2.6.2.4 Usability testing settings* below for more details related to how testing is carried out).

Preference data are often solicited during post-test debriefing, such as when users are asked to complete questionnaires or to participate in follow-up interviews (Nielsen, 1993). During debriefing sessions, users may be asked to comment on the ease of use and learnability of the product, how well the product allowed them to meet their goals, their likes/dislikes, and/or suggestions for improvements (Nielsen, 1993; Rubin & Chisnell, 2008). Verbal and non-verbal data obtained during the testing session may also be helpful in this regard. For example, the incidences of non-verbal frustration that are noted during a testing session could be recorded (e.g., pounding the keyboard in frustration) as an indication of user dissatisfaction (Hartson, et al., 1996).

Another useful source of usability data are verbalizations made by users during the testing session while using the "think aloud" protocol which is described in detail below.

2.6.2.3.1 The think aloud protocol. The think aloud protocol is commonly employed in usability testing to allow test administrators to gain insight into what users are thinking as they complete tasks. According to Nielsen, "Thinking aloud may be the single most valuable usability engineering method" (1993, p. 195). For example, if users think aloud and verbalize their thoughts as they complete an assigned task, it could be possible to obtain more information about why a particular problem was encountered and could thus help researchers to better log the incident. Tamler (2001) offers a useful example to illustrate the value of the think aloud protocol: Clicking on the incorrect button could be logged as a user error that the researcher may perceive to occur due to a confusing icon, however, user verbalizations may indicate that the button was clicked out of interest or because it was situated too close to the adjacent button. This example demonstrates that, without access to the thoughts of the user, it is not always possible to understand the basis behind the user's actions. The think aloud protocol may also allow for the collection of preference data if the user verbalizes pleasure or frustration during the test.

While many studies cite the method described by Ericsson and Simon (1993) as the source of think aloud protocols used in usability testing, the literature indicates that such studies do not always adhere to this traditional method and that there are in fact several versions of think aloud techniques currently in use (Boren & Ramey, 2000; Olmsted-Hawala, Murphy, Hawala, & Ashenfelter, 2010a, 2010b). Ericsson and Simon's traditional think aloud protocol stipulates that users' thoughts that arise directly from interaction with the product (and which are not "filtered" or "processed" in order to interpret events or offer opinion about experiences) are the most valuable verbalizations for analysis (Ericsson & Simon, 1993). Moreover, the test administrator should only prompt users to verbalize if a previously specified period of silence has elapsed and it is necessary to remind users to think aloud. However, Boren and Ramey (2000) remind researchers that Ericsson and Simon's think aloud parameters were developed for research in the field of cognitive psychology (for which research goals differ from that of usability testing), and that comments related to the subjective thoughts of the user are often highly valuable in usability testing contexts. As a result, it is not necessary or even always desirable to stringently follow the traditional think aloud technique.

Other think aloud techniques include a coaching method in which the test administrator frequently asks the user direct questions, and a method grounded in speech communication theory in which the test administrator acts as an active listener and frequently offers verbal cues to indicate that he/she is listening and may gently prompt the user to think aloud or even clarify comments as needed (Boren & Ramey, 2000). Olmsted-Hawala and colleagues have found that there was no statistically significant difference in the amount of feedback offered by users when traditional, speech communication, or coaching think aloud techniques were compared (Olmsted-Hawala, et al., 2010b). However, those users that were subject to a coaching method tended to perform the assigned tasks more accurately and expressed more satisfaction with the website being tested (Olmsted-Hawala, et al., 2010a). This suggests that if the purpose of the usability testing is to identify how usable a product will be when the user works without help, the traditional or speech communication versions of the think aloud protocol may be most appropriate.

In addition to considering which think aloud technique is most appropriate, it is also necessary to consider whether this approach is compatible with other parameters of the study, namely the nature of the data to be collected and the characteristics of the users. There is conflicting information available about whether or not employing the think aloud protocol will slow task completion (Lewis, 2006). The prevalent viewpoint is that use of the think aloud protocol will significantly slow users (Hertzum, Hansen, & Andersen, 2009; Nielsen, 1993; Rubin & Chisnell, 2008), however, there are also data that are contrary to this (Olmsted-Hawala, et al., 2010a). In order to err on the side of caution, it may be wise to avoid use of this technique when the efficiency of task completion is a metric that is examined for. Rubin and Chisnell (2008) also caution about the challenges of thinking aloud for users with cognitive disabilities who may have difficulty thinking aloud as they concentrate on completing the assigned task (p. 55).

2.6.2.4 Usability testing settings. There are a variety of settings that may support usability testing, ranging from a traditional testing laboratory, to a portable testing lab, to remote testing. Financial considerations, characteristics of the user population, the desired number of users included in testing, and the purpose of the testing may influence which testing setting is most appropriate.

The traditional testing laboratory includes a test room and may also include a second adjacent observation room (Nielsen, 1993; Rubin & Chisnell, 2008). The test room is where the user (tester), test administrator, and one or more additional observers (e.g., notetakers, if required) reside. There is a workstation for the user with a computer and any other equipment that the user may need. The test administrator is seated behind the user at a 45 degree angle so that he/she is in the peripheral vision of the user but not so visible as to be distracting, and additional observers may be seated behind the user (Rubin & Chisnell, 2008). Cameras may be present to videotape the user's computer screen (if screen recording software is not installed on the user's computer), face, and/or workstation. If there is an adjacent observation room, it is typically separated from the test room by a sound proof wall and one-way mirror, and may include monitors that display the views from video cameras present in the testing room (Nielsen, 1993). Rubin and Chisnell (2008) suggest that one or more individuals with technical expertise about the product being tested are also available to address unexpected problems that may arise that the test administrator may not be able to deal with.

If a traditional testing laboratory is not available and/or it is not practical for the users to visit the laboratory, portable testing may be employed. In the portable testing method, the test administrator visits the user (for example, at the workplace of the user) with equipment and supplies required to complete the test. Testing equipment and supplies may include a laptop (e.g., with an application to be tested installed, or access to a website under study), a means of recording the session (e.g., video camera(s) and tripod, microphones, and/or screen recording software installed on the laptop), and forms (e.g., consent forms, questionnaires) (Nielsen, 1993; Rubin & Chisnell, 2008). The test administrator may set up the test room in an office at the workplace of the user. A variation of this method is the use of usability kiosks, whereby a "self-serve" testing station is set up in a heavily trafficked area in an office and users that pass by the kiosk may stop to try out a user interface on display and offer feedback (Nielsen, 1993).

While traditional lab testing allows the test administrator a large degree of control over the testing circumstances (Bartek & Cheatham, 2003a) and portable testing allows for testing to take place in a realistic environment in a manner that is convenient to the user, a major disadvantage of these methods is the financial and time commitments associated with set-up and travel carried out by the test administrator and/or user. Remote testing, which Hartson and colleagues (1996) have defined as "usability evaluation wherein the evaluator, performing observation and analysis, is separated in space and/or time from the user" (p. 228) offers unique advantages that may alleviate some of the drawbacks of traditional and portable laboratory testing. Depending on the availability of the product to be tested (e.g., an open source application available on the Internet, or licensed software installed on computers in the user's workplace) the user may participate in the testing from various locations including home or work. Remote usability testing is thus more convenient for users who are busy or physically unable to attend the laboratory (e.g., due to illness, disability, or geographic location) (Baravalle & Lanfranchi, 2003; Houck-Whitaker, 2005) or for whom laboratory testing is less practical/realistic from testing in the user's own environment – the latter condition arising, for example, when users with disabilities who use assistive technologies to interact with a product are the desired testers (Power, Petrie, & Mitchell, 2009). Remote testing may also be more comfortable for users who find a traditional testing laboratory intimidating. Moreover, extending the potential pool of users worldwide via remote testing is useful when the desired users have highly specialized skills (Bartek & Cheatham, 2003a).

Depending on the nature of interaction between the user and test administrator, remote usability testing is referred to as synchronous (when the user and test administrator interact electronically in real time during the test) or asynchronous (when there is no interaction between users and test administrator while the user completes the test). Early reports of remote usability testing describe synchronous remote testing. For example, Hammontree, Weiler, and Nayak (1994) described a method of usability testing whereby video conferencing software, Web cams, and a telephone line facilitated sharing of the user's computer screen and allowed the user and test administrator to see and speak to each other in real time.

Several published reports comparing traditional usability testing to synchronous remote usability testing have since emerged and have indicated that there is typically no significant difference in the effectiveness of traditional vs. synchronous remote usability testing in terms of the amount of critical incidents identified (for example, refer to Andreason, Villemann Nielsen, Ormholt Schroeder, & Stage, 2007; Hartson, et al., 1996; McFadden, Hager, Elie, & Blackwell, 2002; Selvaraj, 2004). Andreason and colleagues (2007) also found that there was no significant difference in the number of tasks successfully completed or task completion time, though McFadden and colleagues (2002) found that time on task may be greater during remote sessions, perhaps due to reduced network performance (e.g., participants waiting for a slowly loading page, and losing their train of thought). Hartson and colleagues (1996) also compared qualitative data obtained from post-testing questionnaires, and found that there was no significant difference in responses for Likert-scaled questions or in the amount of responses for optional open-ended questions following traditional vs. synchronous remote testing.

When asynchronous remote usability testing is conducted, some or all of the data collection is automated via special usability testing software (which also may be used, to some extent, during traditional usability evaluation). For example, usability testing software may record quantitative data such as mouse clicks, time spent on pages, and task efficiency (Bartek & Cheatham, 2003b; Dray & Siegel, 2004; Kraus, 2003). This may be augmented by completion of online surveys and/or submission of written comments by the user which are usually submitted after completion of the assigned tasks. Because of the automated collection of data and the potential for (at least partially) automated data analysis by usability testing software, asynchronous remote usability testing has emerged as a useful method when a product is likely to used by a large and diverse population of users and thus it is desired to include a larger number of users in the testing (Baravalle & Lanfranchi, 2003; Harty, 2011; Tamler, 2001). Additionally, this method may be valuable for validation (verification) testing whereby quantitative metrics may be compared to previously established usability benchmarks (e.g., desired efficiency of task completion).

There are a few published reports that compare the effectiveness of synchronous and asynchronous usability testing methods. While the amount of quantitative data collected has been found to be comparable across both methods, there may be a significant reduction in the amount and quality of qualitative data collected by asynchronous testing (Petrie, Hamilton, King, & Pavan, 2006). This reduction in qualitative data has been attributed to a reduction in think aloud data from users and lack of user-test administrator interaction in asynchronous tests (Petrie, et al., 2006; Tamler, 2001). Tamler (2001) points to the value of think aloud data from users when they become frustrated, and the role of the test administrator in asking for clarification about comments and helping the user to articulate why they are experiencing frustration. Petrie and colleagues also highlight the value of observing how users who work with assistive technologies interact with the application under study – observations that could only be made in traditional or remote testing when the user's computer screen and/or the user is observed or videotaped (Petrie, et al., 2006).

Due to the impact on data collection as a result of the choice of remote usability method, some researchers recommend supplementing remote testing with traditional methods (Harty, 2011; Tamler, 2001) and to use synchronous methods when qualitative data are desired (e.g., for exploratory/formative testing) and asynchronous methods when large amounts of quantitative data are desired (e.g., for assessment/summative testing in which statistical analysis of data is required) (Dray & Siegel, 2004; Petrie, et al., 2006).

2.6.3 Applying usability testing methods to e-learning accessibility testing. From this review of usability testing, it is clear that the methods described may be useful for accessibility testing of e-learning infrastructure in order to evaluate the accessibility of technologies for anticipated user populations. For example, like usability testing, accessibility testing would benefit from including users (i.e., students) as the testers. Like usability, there are subjective attributes to accessibility such as user satisfaction, and thus collecting user preference data is also relevant with respect to accessibility. In order for e-learning infrastructure to be accessible, it must be usable and thus performance metrics may also be of value in accessibility testing.

Formative, summative, and validation testing methods may also apply to e-learning accessibility testing. For example, exploratory (formative) accessibility testing may be valuable during the early stages of design of a new LMS, assessment (summative) evaluation may be helpful to identify if accessibility problems identified during formative assessment of the system have been addressed by asking students to work through assigned tasks in the system, and validation (verification) testing may be carried out to ensure that automated or expert conformance evaluation (e.g., conformance to WCAG 2.0) does indeed correlate with the student experience. Indeed, validation testing using the WCAG 1.0 as benchmarks was the method of LMS accessibility testing described by Power and colleagues (Power, et al., 2010).

Methods similar to those described for usability testing would thus allow for accessibility testing, and the choice of user characteristics (e.g., a concerted effort to include a diverse student tester population) as well as supplementary data collected (e.g., the choice of questions included in pre- and post-task questionnaires and/or interviews) would serve to further tailor the testing to solicit accessibility data.

2.7 Conclusions and Specific Research Questions

This review of the literature has shown that e-learning accessibility is a multi-faceted and complex issue. The increasing impact of Web accessibility guidelines such as the WCAG on the design of learning technologies and e-learning environments is a helpful step towards enhancing the accessibility of e-learning in higher education. However, the highly subjective nature of accessibility, coupled with concerns that have been expressed about the limitations of accessibility guidelines, strongly points to the need to augment objective accessibility testing with subjective feedback. In the context of e-learning accessibility in higher education, feedback from students would be most relevant and it is likely that students with disabilities would have much to offer with respect to identification of disabling attributes of e-learning environments. As a result, it is important to explore how objective evaluation of the accessibility of e-learning infrastructure compares to subjective evaluation of accessibility by students (including students with disabilities). To this end, the first specific research objective of this project was to:

1. Determine the extent to which objective measures of the accessibility of e-learning technologies are able to predict the subjective accessibility experience of students.

This literature review has also highlighted the fact that accessibility and usability are very closely related, if not overlapping, attributes. As such, borrowing methods from the well-established field of usability testing to be applied towards e-learning accessibility testing is a valid approach. As with usability testing, it is essential that the methods chosen to conduct accessibility testing take into account characteristics and needs of the user population and, in this context, the users are post-secondary students with a wide range of learning needs and preferences. There are several methods of usability testing that could be adapted for e-learning accessibility testing with student users and it is necessary to compare the suitability of different approaches. As a result, the second research objective of this project was directed towards exploring methodologies for subjective e-learning accessibility evaluation. Specifically, the second research objective of this project was to:

2. Determine whether data obtained from moderated and unmoderated e-learning accessibility testing are different and, if so, how and why they differ.

CHAPTER 3: Methodology

This dissertation presents an empirical study with a mixed methods research design. The choice of methods of data collection and analyses were based on the research questions, research methodology used in the field of usability testing, and the intended audience of this study.

This study compares data about the accessibility of e-learning infrastructure that has been obtained from objective examination (i.e., automated analyses) and from subjective examination (i.e., from student testers), while also comparing two different methods of subjective accessibility testing. Quantitative analysis of numerical data by statistical methods has been conducted to determine whether there are statistically significant differences in the data obtained from the different formats of subjective accessibility testing. This type of analysis dominates in the related field of usability testing and, in this context, may be deemed particularly credible by certain stakeholders such as educational administrators and vendors of educational technologies who may have a preference for numerical data that are somewhat generalizable. At the same time, the focus of this study is accessibility – a highly subjective variable for which qualitative analysis of insight offered by student participants is expected to be very relevant. For example, in order to understand why a student may prefer to participate in one method of accessibility testing versus another, or how a student may have felt enabled or disabled by educational technology (by identifying enabling or disabling features of e-learning infrastructure), analysis of qualitative data allows for development of more in-depth insight into questions such as these than would be possible from statistical hypothesis testing alone. Moreover, themes revealed from qualitative analysis of student data further aid in an understanding of why particular trends were observed from statistical analyses and may help to provide a clearer picture of the experiences of students as they interact with e-learning infrastructure and participate in accessibility testing.

Prior to a discussion of specific methods that were utilized, it is necessary to first consider how the construct of disability has been defined in this study and how that definition has influenced the way in which e-learning accessibility has been defined and examined in this study.

3.1 Defining Digital Disability and E-Learning Accessibility

In this study, theory has informed the manner in which disability and e-learning accessibility have been defined and has, as a result, influenced both the methods of data collection and data analysis.

Digital disability has been considered in a manner that is aligned with the social model of disability. This model of disability (as described in Marks, 1997; Oliver, 1996) differs from the prevalent medical model of disability in the way that impairment and disability are defined relative to one another. The medical model suggests that impairments are responsible for disabilities. For example, loss of vision leads to a print disability when a blind student wishes to read a text that is not available in Braille format. In contrast, the social model of disability suggests that impairment and disability are distinct because impairment leads to disability only when socially-constructed barriers are encountered (Oliver, 1996). Applying this model to the same example, it is not the visual impairment that is responsible for the disability (the inability of the student to read the book) but rather the socially-constructed norm of practice followed by the publisher of the book in choosing not to make a Braille version available. There are criticisms of this model, including the manner in which impairment remains rooted in medical terms and the removal of the corporeal (bodily) aspect of disability such as emotional consequences, pain, and fatigue (Hughes & Paterson, 1997; Terzi, 2004). However, in spite of these limitations, the social model of disability is useful in this context because it prompts consideration of the possibility that disabling features representative of social norms may be embedded into the design of learning technologies.

The idea of social construction of technology is not a new one. In her 1989 Massey Lectures, Ursula Franklin asserted that technology is a practice and not simply a tool because it provides users with ways of doing things (Franklin, 2004). If technology is deemed a practice, and practice is considered to be a representation of socially-accepted cultural norms, technological design can therefore be seen as a reflection of those norms. Contemporary scholars in critical disability studies have also suggested that socially-defined norms of use are routinely embedded into the design of technologies. For example, Goggin and Newell have suggested that there has not been sufficient examination of the "disablism" that is represented in some technologies when disability is built into the technology in the form of non-disabled norms of use (Goggin & Newell, 2003; Newell, 2008).

Considering digital disability from the vantage point of the social model may raise the profile of e-learning accessibility because it supports the notion that all students (regardless of whether or not they deem themselves to be disabled in the traditional medical model sense) could potentially be enabled or disabled by characteristics of learning technologies. Another related model that is relevant to this discussion is the biopsychosocial model of disability. In describing the International Classification of Functioning, Disability and Health (ICF), a framework used by the WHO to measure non-fatal health outcomes, Üstün and colleagues (2003) introduce the notion of a biopsychosocial model as an intermediary model that recognizes the interaction between characteristics inherent to an individual (medical factors) and external (social) factors as contributors to disability. To understand how these models of disability can aid in a conceptualization of disability in a digital context, consider the following example: While a student with a physical disability who experiences difficulty using a mouse could be disabled by the design of a keyboard-inaccessible website, so too could a student who does not identify him/herself as disabled but who finds a website inaccessible due to the need for a hi-speed Internet connection. Therefore accessibility of e-learning infrastructure is relevant to all students.

Digital disability was defined in this study as the inability of a student to interact with e-learning infrastructure in a way that meets his or her needs in a particular context as a result of properties that are inherent in the design of the technology. In order to explicate this statement, recall the definition of accessibility offered by the IMS Global Learning Consortium that was first discussed in Chapter 2:

> Accessibility…is the ability of the learning environment to adjust to the needs of all learners…. The needs and preferences of a user may arise from the context or environment the user is in, the tools available…, their background, or a disability in the traditional [medical model] sense. Accessible systems adjust the user interface of the learning environment, locate needed resources and adjust the properties of the resources to match the needs and preferences of the user. (IMS Global Learning Consortium, 2004)

This statement suggests that the burden of adaptability lies with the learning environment rather than with the student, and the learning environment may thus enable or disable the student. Accessible e-learning environments must therefore be flexible and suitable for use by diverse populations of students with a variety of learning needs and preferences (shaped by their ability or disability as well as learning context). This notion is also aligned with Vanderheiden's description of universal usability as "a focus on designing products so that they are usable by the widest range of people operating in the widest range of situations as is commercially practical" (Vanderheiden, 2000, p. 32). As such, the theoretical framework of this study supports the notion of accessibility and usability as inter-related (rather than distinct) attributes.

In addition to informing the definitions of digital disability and e-learning accessibility applied in this study, and supporting the inclusion of a diverse collection of student participants, the theoretical framework was also applied to the manner in which data were analyzed. Data collected in this study were analyzed with the goal of identifying enabling or disabling features of the e-learning infrastructure rather than the impact of student impairment on accessibility. For example, data obtained was categorized according to learning technology (e.g., accessible and inaccessible features of a PowerPoint presentation) rather than student characteristics (e.g., problems experienced by students with disabilities vs. problems experienced by non-disabled students). While it is likely that students with different characteristics will experience different benefits and challenges, it is also possible that removal of barriers identified by one student will also benefit another student who could experience the same barrier in another context. For example, a non-disabled student who finds it necessary to study in a dimly lit location may benefit from flexible e-learning attributes that also routinely benefit students with print disabilities. As such, it is not necessarily important to describe accessibility in terms of specific abilities or disabilities. It is, however, useful to identify general benefits and challenges associated with e-learning in order to improve e-learning accessibility for all students. The goal of this research program was to aid in the development of methods of e-learning accessibility testing in order to accomplish just that.

3.2 Student Recruitment

A total of 24 students were recruited for this study ($N = 12$ students with one or more learning disability, and $N = 12$ students who did not identify as persons with disabilities). The rationale for including students with learning disabilities in the study was three-fold. Firstly, recruiting students with one type of disability (vs. a wider variety of disabilities) was intended to reduce the complexity of the study. Secondly, feedback from students with learning disabilities in particular was deemed to be particularly relevant in this context, given the special attention that has appeared in the literature related to the limitations of Web accessibility guidelines in meeting the needs of users with learning disabilities. Thirdly, students with learning disabilities may experience barriers that affect a wide array of academically-relevant skills, including listening, understanding written and oral instruction, and written expression – areas in which all students may struggle to varying degrees in different contexts. Therefore inclusion of students with learning disabilities and students who do not identify as students with disabilities in this study allowed for recruitment of a group of participants with a wide array of learning needs and preferences, who may be expected to encounter similar barriers and facilitators to accessibility, albeit to varying degrees. This is aligned with the way in which disability has been conceptualized in this study, and an interest in focusing on attributes of the learning environment rather than ability/disability status of the students when analyzing results of the study.

Several simultaneous efforts were made to recruit student participants. Flyers advertising the study were posted on campus (refer to Appendix A for a sample recruitment flyer), and a recruitment email (including the same language as the flyer) was sent to students registered with the Learning Disabilities division of the campus Counselling & Disability Services office. The snowball sampling technique, which involves asking participants to refer others who may be interested in participating, was also employed. Screening interviews were conducted by email to determine whether interested participants identify as persons with disabilities (to ensure that an equal number of students with and without disabilities was recruited), to learn about typical challenges that the students face when working on the computer, and to ensure that it would be possible to meet the needs of students with respect to access to the testing laboratory and computer (refer to Appendix B for sample screening interview questions). Students who were not very forthcoming during the screening interview stage were not pressed for additional personal details, besides gently explaining why it was necessary for the researchers to know whether or not the student identified with a learning disability.

3.3 E-Learning Infrastructure Tested

Two e-learning scenarios were prepared so that each student participant could complete both, one in an unmoderated session and the other in a moderated session (refer to section *3.5 Accessibility Testing with Student Users* for more details about how the testing sessions were organized). Each module consisted of a mock online course module. Students were asked to access the modules from a mock online course hosted within Moodle 2.0. While Moodle is the LMS in use at York University, the online course was hosted within the Moodle server of the research unit ABEL (Advanced Broadband Enabled Learning), and therefore the appearance of the course did not adhere to the standard appearance of York University Moodle pages that some of the students had previous experience with. A course shell for a mock course called "Introduction to Digital Literacy" was created and populated with content that students accessed via dummy student accounts created for this study. Within the *Introduction to Digital Literacy* course, two modules titled *Scholarly Resources* and *Credible Resources* were created.

The modules were designed to closely match each other in format, style of content, and anticipated accessibility of webpages included within the modules (as characterized by the number of potential accessibility problems identified by objective accessibility measures as described in section *3.4 Objective Accessibility Testing*). The modules were designed with Nielsen's (1993) advice to include realistic tasks that leave users feeling a sense of accomplishment upon completion in mind. Each module consisted of a single page that included instructional content in the form of an embedded PowerPoint presentation, an embedded YouTube video, and a task to be completed that required students to access websites external to Moodle. At the end of each module homepage was a link to a discussion forum where students finished the module by commenting on their experience in completing the task. The Scholarly Resources module introduced students to the definition of scholarly resources and required them to locate a scholarly article, while the Credible Resources module introduced students to methods that can be used to evaluate online resources, and required students to evaluate two different webpages. Accessibility barriers were not deliberately included within the modules. Refer to Appendices C and D for details and instructional materials relevant to the Scholarly Resources and Credible Resources modules, respectively.

3.4 Objective Accessibility Testing

Objective methods were used to assess the accessibility of e-learning infrastructure included in each module of the mock online course. To do this, a set of open source online tools were used for automated accessibility testing of websites and mock course materials as follows: AChecker (http://achecker.ca/checker/index.php), the Qompliance add-on for Firefox (https://addons.mozilla.org/en-US/firefox/addon/qompliance/), and the WAVE web accessibility evaluation tool (http://wave.webaim.org/) were used to examine websites that all students accessed for potential accessibility problems. The accessibility evaluation feature of PowerPoint 2010 was used to examine the module PowerPoint files for potential accessibility problems, and Power Talk (http://fullmeasure.co.uk/powertalk/), an open source screen reader for PowerPoint files, was used to verify the output from the PowerPoint accessibility evaluation. The accessibility data collected for all components of the e-learning infrastructure were tabulated to allow for comparison with data obtained from student accessibility testing, described next.

3.5 Accessibility Testing with Student Users

A 2 x 2 counterbalanced within-subject testing design was employed. The 2 x 2 designation refers to the inclusion of 2 factors, each with 2 levels. The factors were course module and method of accessibility testing, and the levels were Module A and Module B (for the Scholarly Resources and Credible Resources modules) and moderated and unmoderated (for the method of testing).

The within-subject approach for the first factor (the module) allowed for the comparison of observations of a particular student across the learning modules. The within-subject design for the second factor (the format of testing session) allowed for comparison of data related to the subjective experience of each student across the two testing methods. The counterbalanced design in which the order that each module was completed varied amongst testing groups was intended to minimize the effect of learning that may be carried over from one module to the next. It is also possible that performance and satisfaction with the testing process may increase as students complete the second task (if they become more comfortable with the testing setting) or decrease (if students feel a growing sense of frustration), and thus the order of testing methods used (moderated vs. unmoderated) was also varied. Students were assigned to testing groups such that there were an equal number of students with and without disabilities so that each testing group could be expected to be approximately equally heterogeneous. Table 1 presents a testing matrix.

Table 1.

Counterbalanced Within-Subject Accessibility Testing Matrix

Testing Group[a]	Test 1	Test 2
1	Scholarly Resources (unmoderated)	Credible Resources (moderated)
2	Scholarly Resources (moderated)	Credible Resources (unmoderated)
3	Credible Resources (unmoderated)	Scholarly Resources (moderated)
4	Credible Resources (moderated)	Scholarly Resources (unmoderated)

[a] Students were assigned to a testing group to determine the order of module and accessibility testing method administered. The testing itself was conducted individually (one student per testing session).

Each student participant attended a single one hour testing session, in which Test 1 and Test 2 (as described in Table 1) took place. By completing both tests in the same session, students were exposed to both testing methods on the same day and were perhaps best able to compare their experiences with each testing method (without, for example, having forgotten how the Test 1 experience compared to the Test 2 experience).

Comparing tests completed in a single session also ensured that other conditions that are not constant from day-to-day yet which could influence the study results will be reasonably constant for both tests (e.g., the degree of cognitive impairment that a student experiences). Table 2 lists the components of a testing session, including an estimate of the approximate time that each component took, on average.

Table 2.

Components of E-learning Accessibility Testing Sessions

Event	Description	Approx. Time
Introduction	A verbal introduction was read to the student, and informed consent was obtained.	5 min
Think Aloud Practice	The think aloud protocol was described and demonstrated, and the student practiced the protocol.	5 min
Test 1	The student completed Test 1 (assigned according to the Test Matrix in Table 1).[a]	20 min
Test 2	The student completed Test 2 (assigned according to the Test Matrix in Table 1).	20 min
Exit Interview	A semi-structured interview was conducted to learn more from the student about their subjective experience in completing the modules and thoughts on both testing formats.	10 min

[a] Students were asked if they would like to take a break between completing Test 1 and Test 2. All but one of the 24 students declined a break.

Refer to Appendices E, F, and G for verbal scripts for the Introduction, Think Aloud practice, and Exit Interview protocol, respectively.

3.5.1 Moderated and unmoderated testing sessions. The testing location utilized for both formats of testing sessions was a classroom that included a work-station for the student, equipped with a computer and a wired Internet connection. At the start of each testing session (moderated and unmoderated), the test administrator assisted the student in launching the data collection software on the computer. Following this, there were differences in the way in which the moderated and unmoderated testing sessions took place.

The unmoderated testing sessions were intended to simulate asynchronous remote testing. Data were collected electronically (online) without the use of video cameras or interaction with the test administrator. Once the testing software was launched on the workstation computer, the test administrator moved to an adjacent observation room that was separated from the testing room via a one-way mirror. Students worked alone in the testing room but were instructed to request the test administrator by signalling via the one-way mirror if they encountered an incident from which they did not know how to proceed (e.g., a major technical issue such as loss of Internet connection, or a problem with the testing software).

Moderated testing sessions were intended to reflect traditional laboratory testing methods. These testing sessions were similar to that described for the unmoderated testing sessions (and took place at the same workstation). Two differences in this format were that: 1) a single video camera in the room recorded the student (including their face and physical interactions with the computer) to capture additional non-verbal visual data; and 2) the test administrator was present in the room and interacted with the student to remind him/her to think aloud as needed, following a speech communication form of the think aloud technique (as previously described in section *2.6.2.3.1 The think aloud protocol*). This version of the think aloud technique was used to maximize the amount of subjective data obtained, without intervening as much as would take place during the coaching method of the technique (the coaching method being less applicable, as students are likely to be accessing e-learning infrastructure on their own in a natural setting). In instances when students requested help from the test administrator, they were instructed to do what they would normally do if they were working independently.

3.5.1.1 OpenVULab. OpenVULab (http://openvulab.org), a Web-based open source screen and verbalization recording tool, was used as the data collection software in this study. OpenVULab is a new tool to support accessibility testing research and had not been used in the current form prior to this study. Therefore, a small pilot study including students with learning disabilities as testers ($N = 7$) was conducted prior to completing this study to ensure that OpenVULab would be accessible to the student participants in this study. The pilot study revealed that the user interface of OpenVULab was highly accessible for students included in the pilot and allowed the development team to make improvements to the tool prior to the start of this study.

Testing sessions began with OpenVULab prompting the student to complete a pre-test survey including multiple choice and yes/no questions aimed at determining whether the student considers him/herself to have one or more disabilities, self-reported skills as relevant to completion of the module, and the degree of prior experience with the technologies relevant to the module (refer to Appendix H for the pre-test questions that were administered). Following completion of the pre-test survey, OpenVULab presented a link to the module homepage within Moodle. As students worked through the relevant module, OpenVULab created a video recording of all on-screen interactions. At the same time, think aloud verbalizations made by the student were also recorded and synced with the video recording. This feature made OpenVULab a particularly useful tool in this context because the lack of think aloud data has been reported as a limitation of asynchronous testing that may result in a reduced amount of subjective qualitative data (Tamler, 2001). Because accessibility is a highly subjective attribute, it was hypothesized that qualitative think aloud data might be highly valuable in this context.

Overall then, OpenVULab created a single audio and video file analogous to the data that would be collected in a traditional laboratory testing session. When the student completed a module, OpenVULab also collected additional data via a post-test questionnaire. The post-test questionnaire questions were Likert-scaled questions aimed at identifying the students' perceptions regarding ease of task completion, accessibility of e-learning infrastructure encountered, and ease and comfort level with participation in the testing method (refer to Appendix I for post-task questions that were administered).

3.6 Data Analysis

Data collection involved viewing OpenVULab recordings of moderated and unmoderated accessibility testing sessions, viewing camera recordings from moderated accessibility testing sessions, reviewing responses to OpenVULab-administered questionnaires, and coding transcripts from exit interviews. Data collected from student testing sessions was also compared to data from objective testing. Refer to Table 3 for an inventory of the data from each of the accessibility testing methods.

3.6.1 Video analysis. Video data (from OpenVULab and video cameras) was analyzed using Transana (http://transana.org/), an open access data analysis and management software program.

OpenVULab screencasts created while students completed each module were reviewed to determine the time required for module completion (efficiency, in minutes), whether or not modules were successfully completed (assessed by determining whether students successfully completed all aspects of a given module), the number of verbal expressions of pleasure or frustration expressed during each module (from think aloud audio), and the number of accessibility challenges (critical incidents) that students encountered while completing each module. Critical incidents were defined as challenges that a student encountered as related to accessing content (e.g., playing a YouTube video), performing a task (e.g., successfully posting a comment to a discussion forum), or accessing an understanding of content (e.g., understanding written or verbal instructions), the identification of which was often aided by relevant think aloud audio. Additionally, review of the OpenVULab screencasts facilitated creating a list of webpages accessed by all students to test for accessibility by objective methods with the tools listed in Table 3. Video camera recordings from moderated testing sessions were also reviewed using Transana. Additional visual and non-verbal information noted included expressions of frustration (e.g., facial expressions and/or frustrated body language) and relevant student-computer interaction (e.g., leaning forward to read small text on the computer monitor).

Table 3.

Inventory of Data Collected from Subjective and Objective Accessibility Testing

Data Collected	Testing Method[a]		
	M	U	O
OpenVULab recordings			
Counts of accessibility problems encountered with all technology	+	+	
Counts of verbal expressions of frustration	+	+	
Counts of verbal expressions of pleasure	+	+	
Efficiency (time to complete the module, in min)	+	+	
Video camera recordings			
Counts of non-verbal expressions of frustration	+		
Counts of non-verbal expressions of pleasure	+		
Descriptions of student-computer interaction	+		
Pre- and post-test OpenVULab questionnaires			
Demographics related to disability and technological experience	+	+	
Perceived ease of working with technology and completing the module	+	+	
Ease and comfort with participation in testing session	+	+	
Exit interview			
Most challenging aspect of completing the module	+	+	
Least challenging aspect of completing the module	+	+	
Most positive aspect of participating in the testing method	+	+	
Least positive aspect of participating in the testing method	+	+	
Enabling or disabling impact of technology	+	+	
AChecker , Qompliance, and WAVE conformance testing[b]			
Lists of potential accessibility problems with each webpage accessed			+
PowerPoint accessibility testing tool			
Lists of potential accessibility problems with each slide			+
PowerTalk screen reader simulation			
Confirmation of accessibility problems identified by PowerPoint			+

[a] M refers to moderated, U refers to unmoderated, and O refers to objective accessibility testing.

[b] Potential problems predicted by these methods were merged to create a single list of potential accessibility problems relevant to each webpage.

[+] Indicates that the data was collected from the relevant testing method.

To facilitate the collection of the data described above, each video was reviewed several times. An initial continuous viewing of each video was conducted to make note of general impressions and areas to review in more detail. During the initial stage of data analysis, this preliminary viewing of the videos led to development of a list of critical incidents, which was reviewed with a second researcher prior to applying codes to any of the videos. A second viewing of all videos was conducted to create clips of relevant video segments and to add codes to the clips based on the previously agreed-upon coding scheme. A third viewing of all coded clips was conducted to review the coding process for consistency.

A method of coding similar to that described by Olmsted-Hawala and colleagues (2010a, 2010b) was used whereby "ci" = critical incident, "vp" = verbal expression of pleasure, "np" = non-verbal expression of pleasure, "vf" = verbal expression of frustration, "nf" = non-verbal expression of frustration, and "sc" = student-computer interaction. Each OpenVULab recording was also coded as "ss" (success) or "ff" (failure) to indicate whether the module was successfully completed. Completion time (efficiency) was recorded in instances where the "ss" code was applied.

The Transana software allows for the application of multiple codes to a single clip, and includes a search engine feature that allows for sorting of clips in various ways according to codes of interest. As such, additional codes were added to each clip to indicate which test and e-learning infrastructure it applied to. To tag clips according to the test, the following codes were added: "su" = Scholarly Resources module (unmoderated), "sm" = Scholarly Resources module (moderated), "cu" = Credible Resources module (unmoderated), and "cm" = Credible Resources module (moderated). To tag clips according to the e-learning infrastructure, the following codes were added: "mo" = Moodle, "pt" = PowerPoint, "yt" = YouTube, "vl" = VCU libraries homepage, "al" = *Academic Leadership* journal homepage, "vr" = VCU research news page, and "ya" = Yahoo! discussion forum (the latter four webpages being pages external to Moodle that students were required to access). For example, if a student encountered difficulty understanding advanced terminology within a PowerPoint slide while completing the Scholarly Resources module in an unmoderated session, the clip that shows this would be coded as "ci," "pt," and "su." These additional layers of coding facilitated subsequent organization of data according to the test and/or the technology, which was useful for subsequent comparison analyses.

3.6.2 Questionnaire analysis. Questionnaire data were retrieved from OpenVULab and tabulated within SPSS. The following descriptive statistics were determined for each question: the mean, median, mode, and range of responses. The pre-session questionnaire data were examined as a whole (i.e., including responses from all participants) as well as by Mann-Whitney U testing ($\alpha = 0.05$) to compare responses from students with and without disabilities to determine whether there were differences in self-reported data across these two groups. The post-session questionnaire data were examined according to the relevant module, and further analyzed by Wilcoxon signed rank testing ($\alpha = 0.05$) to determine if the responses for each module differed significantly according to the testing format (moderated vs. unmoderated). In instances where multiple univariate comparisons were performed, the Bonferroni correction was applied to adjust the α in order to reduce the likelihood of inflated Type I error.

3.6.3 Interview analysis. Semi-structured exit interviews with students were recorded and the digital audio recordings were reviewed using Sony's Digital Voice Editor 3 software (available from http://esupport.sony.com/). An initial review of each recording was conducted in order to make notes about possible themes. Each recording was reviewed again, and a transcript was prepared within Microsoft Word. Transcripts were read and open coded to identify themes regarding students' experiences in completing the modules, participating in the study, participating in the two different testing methods, and the impact that technology may have on accessibility of online learning. Relevant quotes from the interviews were then categorized according to the themes of familiarity, simplicity, engagement and learning style, learner control, tolerance, and social presence within a Microsoft Excel document.

3.6.4 Data analysis relevant to Research Objective 1. Recall that Research Objective 1 was to determine the extent to which objective measures of the accessibility of e-learning technologies are able to predict the subjective accessibility experience of students. Two approaches were taken to achieve this, namely, determination of the precision of objective testing in predicting subjective testing results, and summarizing and considering the usefulness of additional qualitative data obtained from subjective testing.

3.6.4.1 Determining the precision of objective data. In order to define the precision of the objective testing, the data obtained from student accessibility testing was considered to be the "true" data, and the objective testing outcomes were compared against this. To accomplish this, critical incidents encountered by students as they interacted with individual pieces of technology were tabulated and prepared for comparison to potential problems identified for the same technologies by the use of automated tools (AChecker, Qompliance, WAVE, PowerPoint, and PowerTalk).

Incorrect predictions about accessibility obtained from objective accessibility testing may be classified as false positive (if an issue was predicted but not experienced) or false negative (if an issue was not predicted, but was experienced). Results concerning false positives may not be useful in this study, since issues predicted by objective measures may not arise given a relatively small group of subjects (as was the case in this study, $N = 24$). A larger and more diverse participant pool may be required for all predicted problems to actually be experienced by users. As a result, only false negative results were examined for in this study. The numbers of false negative results were determined for each webpage that all students encountered, as an indication of deficiencies in the objective testing methods to identify potential accessibility problems. The false negative data were further characterized by examining the nature of the false negative issues to determine if there are particular types of critical incidents that are more readily detected by subjective vs. objective testing.

3.6.4.2 Examining additional qualitative data from subjective testing. Additional data obtained from subjective testing that were not utilized in the precision testing described above was collated. These data include sentiment displays noted in videos (i.e., verbal and non-verbal expressions of frustration, verbal expressions of pleasure, and student-computer interaction), student perceptions of the ease of working with different types of technologies (from post-test questionnaires; refer to Appendix I), and themes revealed by exit interviews (refer to Appendix G). These data were examined to identify whether or not they may shed more light on the degree of accessibility of the technology as well as the perceived impact of the technology in terms of disabling or enabling the students.

3.6.4.3 Summary. Figure 1 presents a schematic overview of this data analysis. Overall, the extent to which false negative results of accessibility problem predictions occur and the extent to which additional subjective data are deemed valuable were applied towards addressing this research objective.

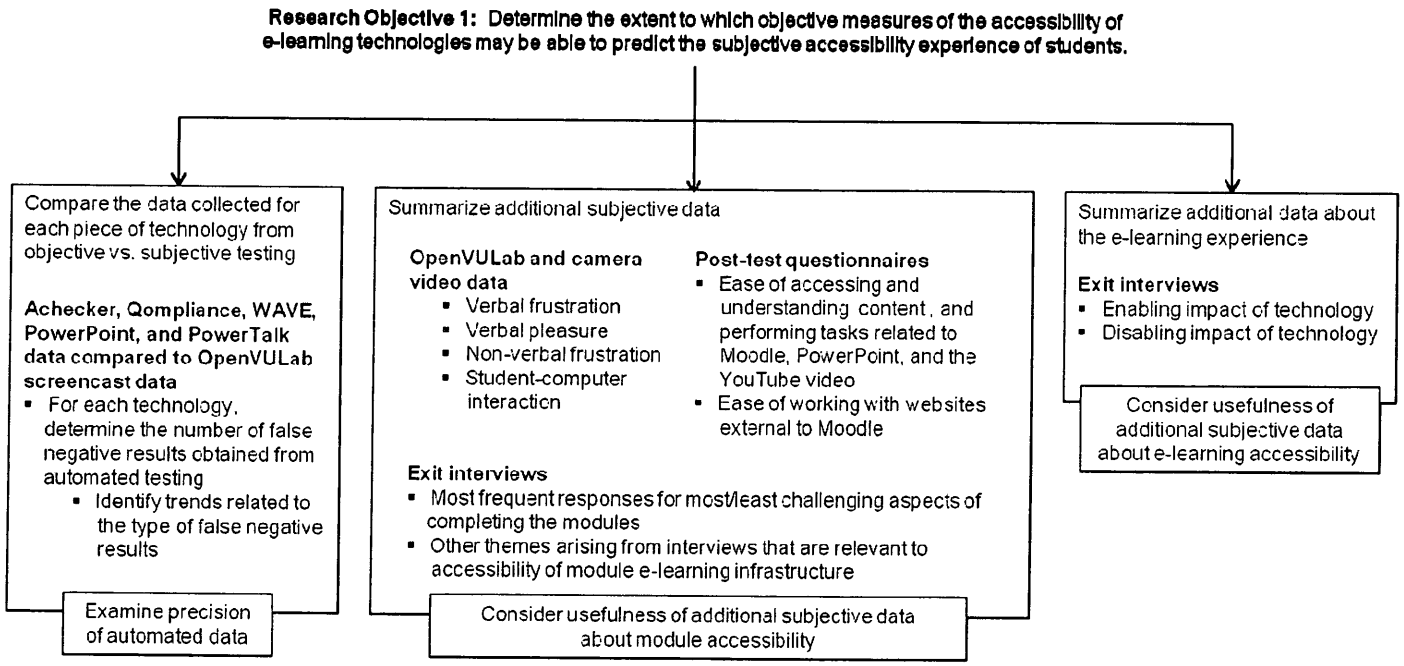


Figure 1. Summary of data analysis relevant to Research Objective 1.

3.6.5 Data analysis relevant to Research Objective 2. Recall that Research Objective 2 was to determine whether data obtained from moderated and unmoderated e-learning accessibility testing are different and, if so, how and why they differ. Three approaches were taken to achieve this, namely, statistical analyses to compare quantitative data obtained from coding of OpenVULab videos created from different testing methods, comparisons of student perceptions of accessibility and participation in the two testing methods (from post-test questionnaires and exit interviews), and description of additional subjective data obtained from video camera recordings and relevant exit interview questions.

3.6.5.1 Statistical analyses of quantitative data. The number of OpenVULab screencast segments coded as "ci" (critical incident), the number of segments coded as "vf" (verbal frustration), and the efficiency of module completion (expressed in minutes) were compared between the moderated and unmoderated accessibility testing methods. To accomplish this, an SPSS worksheet was created for each of these three variables. Within each worksheet, data were tabulated across all of the conditions (i.e., for each module and testing method combination). Statistical analyses were then conducted using SPSS in which both the testing method and module were assigned as independent variables (factors) that may impact one or more of the response variables (ci, vf, and efficiency). This analysis approach allowed for comparison of the response variables from the two testing methods, while also identifying whether the module itself or the combination of module and testing method significantly affected the response variable.

Initial exploratory analysis of each data set was conducted using SPSS to determine which statistical analysis was most appropriate for each response variable. As there were only two instances of "vp," verbal pleasure, noted across all of the screencasts, the "vp" data were not subjected to statistical analysis. Following consultation with the Statistical Consultation Services of York University's Institute for Social Research, the efficiency, critical incident, and verbal frustration data were analyzed by different statistical approaches. The efficiency data were found to most closely exhibit a normal distribution and were examined by two-way ANOVA analysis to determine whether the time it took students to complete each module was significantly affected by the testing method and/or module itself. The critical incident data were found to most closely exhibit a Poisson distribution (Upton & Cook, 2008a) and were examined via Poisson regression analysis (Upton & Cook, 2008b) to determine whether the experimental conditions (module and/or format) significantly influenced the critical incident counts. The verbal frustration data were examined by various methods of regression analysis in an effort to determine whether the testing format and/or module had a significant impact on the counts of verbal frustration. Additional details regarding the rationale for performing these specific analyses are presented in the *Results* chapter.

3.6.5.2 Comparisons of student perceptions. Post-test questionnaire data (refer to Appendix I) obtained following both methods of testing were examined and prepared for comparison. These questions provided data about student perceptions regarding ease of working with different learning technologies, ease of completion of each module, and ease and comfort level associated with completing the testing session in each format.

Themes from exit interview questions relevant to the testing format were also examined (refer to questions 3 to 7 in Appendix G).

3.6.5.3 Description of additional subjective data. Data captured from video cameras and exit interview questions related to most and least challenging aspects of module completion (refer to questions 1 and 2 in Appendix G) were examined. These data were summarized and qualitatively evaluated with respect to whether or not they may enhance the quality of other data (namely, OpenVULab videos and post-test questionnaires) such as by offering more insight into the impact of problems encountered and/or leading to better understanding of why particular problems were encountered.

3.6.5.4 Summary. Figure 2 presents a schematic overview of this data analysis. Overall, comparison of data obtained from moderated and unmoderated accessibility testing included statistical analyses of response variables collected from both methods, comparison of student perceptions of accessibility and participating in the two testing methods, and examination of the value of additional data collected from the video camera and exit interviews.

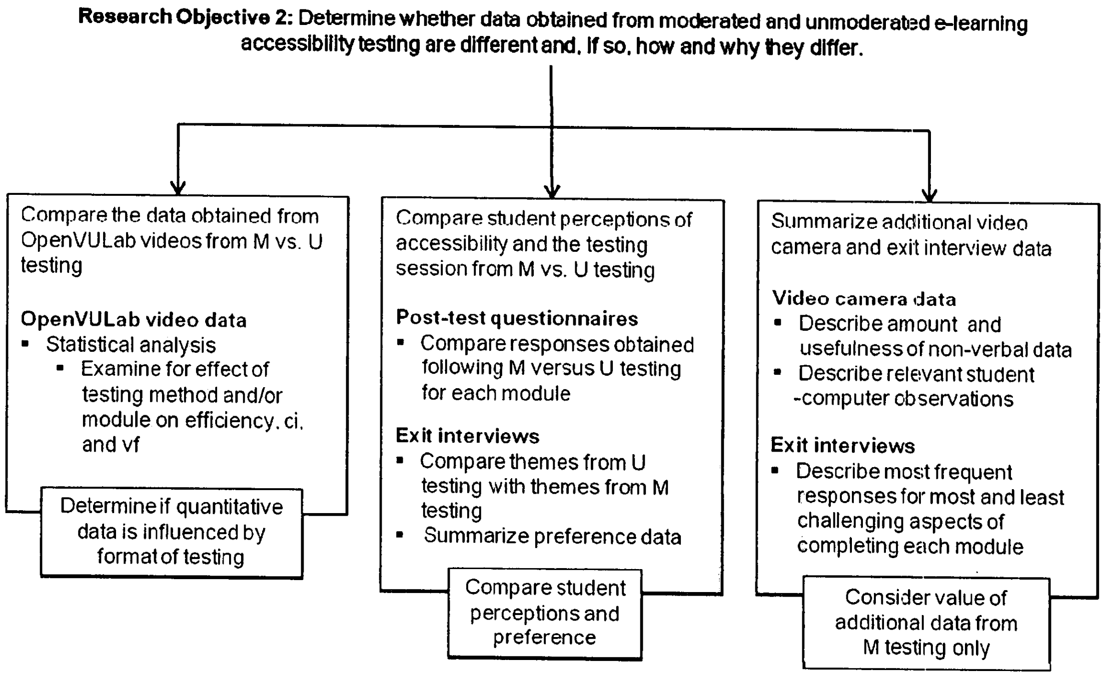


Figure 2. Summary of data analysis relevant to Research Objective 2. M = moderated accessibility testing; U = unmoderated accessibility testing; ci = critical incident; vf = verbal expressions of frustration.

3.7 Summary of Methodology

In summary, in order to determine the extent to which objective measures of accessibility of e-learning technologies are able to predict the subjective accessibility experience of students (Research Objective 1), data from accessibility testing sessions with students and data obtained from objective measures of accessibility evaluation were collected and prepared for comparison. In order to facilitate comparison of data obtained from moderated and unmoderated e-learning accessibility testing (Research Objective 2), data from moderated and unmoderated testing in which the testing conditions differed based on the presence or absence of video capture of the student's interactions with the computer and interactions with a test administrator were also prepared for comparison. The mixed methods research design employed in this study allowed for collection of a rich set of data that may be somewhat generalizable yet will also offer in-depth understandings of how individual students experience e-learning infrastructure and participation in e-learning accessibility testing.

CHAPTER 4: RESULTS

4.1 Self-Reported Participant Characteristics

Five male and 19 female students participated in the study. Refer to Table 4 for a participant list including self-reported data provided by students via email during the recruitment process (see Appendix B) and answers to selected pre-test questionnaire questions (see Appendix H). On the questionnaire, three of the male participants and nine of the female participants self-identified as persons with learning disabilities. One of the female participants identified as a person with a mental health disability in addition to a learning disability. One half of the participants reported having taken an online course before (five of the students with disabilities, and seven of the students without disabilities). The questionnaire also gathered data regarding computer skills, other skills relevant to module completion, and past experience with relevant e-learning technologies. Refer to Table 5 for a summary of these pre-test questionnaire responses, including the mean, median, mode, and range for the responses obtained for each question.

Examination of the medians and modes for the responses from all participants revealed that students most commonly rated their computer and other relevant skills as "good," and that past experience with relevant technologies was rated as "moderate" to "good." Examination of the range for each question revealed that the range of responses given by students with disabilities was greater for 11 of the 15 questions when compared to the range of responses from the students without disabilities. This suggests that the participants with disabilities may have exhibited a wider range of skills and experiences compared to the participants without disabilities. Their inclusion in the study may therefore have helped towards including students with a wide array of learning needs and preferences, as was desired.

To explore whether there were statistically significant differences in the pre-test questionnaire responses from the students with and without disabilities, Mann-Whitney U comparisons were performed on individual questions. This non-parametric test was chosen because exploratory analysis using SPSS revealed that the data were not normally distributed and thus the more powerful independent t test was deemed inappropriate. Moreover, given that the data was not truly interval data (i.e., it cannot be known whether students felt that the options for each question were equally spaced), non-parametric testing that examines the medians may be more appropriate (Kuzon, Urbanchek, & McCabe, 1996). This analysis revealed that there was a statistically significant difference in the responses obtained for the questions "How would you rate your ability to understand written instructions?" and "How would you rate your ability to understand verbal instructions" across the two groups of students ($U = 28$, one-tailed $p = 0.005$ and $U = 21$, one-tailed $p = 0.001$, respectively) with the students with learning disabilities rating their abilities lower. Note that one-tailed analyses were carried out for these questions as it was expected that students with learning disabilities may self-rate these abilities lower than students without disabilities. For a visual depiction of how these data differed across the two populations of students, refer also to Figure 3, which presents the data in the form of clustered bar charts.

It is possible that conducting multiple post hoc comparisons by this method may result in some spurious findings (i.e., inflated Type I error, increased likelihood of finding significant results when they do not exist) (Abdi, 2007). When the α is adjusted from 0.05 to 0.0038 by the Bonferroni method (Bland, 2000; Elston & Johnson, 2008) to account for 15 individual comparisons, only the medians of the responses for the question related to understanding verbal instructions ($U = 21, p = 0.001$) would be deemed significantly different.

Table 4.

Participants Included in the Study

Pseudonym	Gender	LD	Prior Online Learning Experience	Common Challenges Experienced When Working Online
Aliya	F	•		Computers that do not allow use of headphones and music to aid in concentration; websites with large blocks of information
Allan	M	•		Difficulty with webpages that require a lot of reading
Allison	F	•	•	Learning how to use new programs
Arla	F			Locating information when webpages are cluttered and links are difficult to locate
Brad	M		•	Locating icons and information on webpages
Carla	F			Links that do not work
Connie	F		•	Encountering webpages that load slowly or are difficult to navigate
David	M	•	•	Poorly organized webpages with hard to locate links
Donna	F	•	•	Working with webpages that are cluttered
Eva	F	•		Locating relevant information or links
Janna	F	•	•	Locating buttons in an online course; accessing streaming videos
Jeremy	M			Encountering pop-ups on websites

Jessica	F		•	Working with webpages with non-intuitive designs
Julie	F			Slow Internet connection
Karen	F	•	•	Being distracted by pop-ups and images on webpages
Leanne	F		•	No frustrations or challenges encountered
Mandy	F	•		Difficulty locating information on webpages
Mia	F		•	No frustrations or challenges encountered
Morgan	F		•	Slow Internet connection; difficulty navigating webpages
Ollie	F	•		Slow Internet connection
Penny	F			Unreliable Internet access; computer malfunction
Peter	M	•		Difficulty with webpages that include lots of information
Shelly	F		•	Webpages that are not user-friendly and hard to navigate; locating information on webpages
Susan	F	•[a]		No information (the student did not respond to screening questions)

Note. Data was self-reported by email during the recruitment process and/or by completion of the pre-test questionnaire. M = male; F = female; LD = self-identified learning disability; • = positive response

[a] Susan also self-identified as a person with a mental health disability.

Table 5.

Participant Characteristics, as Self-Reported in Questionnaires Administered Prior to Completion of E-Learning Modules

Question	Mean			Median			Mode			Range		
	All	ND	LD	All	ND	LD	All	ND	LD	All	ND	LD
How would you rate your overall computer skills?	3.9	4.1	3.8	4	4	4	4	4	3[a]	2	2	2
How would you rate your ability to navigate through websites?	3.6	3.9	3.3	4	4	3.5	4	4	4	3	1	3
How would you rate your ability to read text on websites?	3.9	4.2	3.6	4	4	4	4	4	4	3	2	3
How would you rate your ability to understand written instructions?	3.9	4.4	3.3	4	4.5	3	5	5	3[a]	3	2	3
How would you rate your ability to understand verbal instructions?	3.8	4.2	3.3	4	4	3	4	4	3	2	1	2
How would you rate your ability to understand visual instructions such as demonstrations?	4.2	4.3	4.0	4	4	4	4	4	4	2	2	2
How often have you used Moodle as a course website?	3.8	4.1	3.5	4	4	4	4	4	4	3	2	3
How often have you viewed PowerPoint files for learning purposes?	3.8	3.6	3.9	4	3.5	4	4	3	4	3	3	3

How often have you viewed videos for learning purposes?	3.3	3.4	3.1	3	3	3	4	3	2[a]	3	2	3
How would you rate your ability to express your thoughts in writing?	3.8	3.9	3.7	4	4	4	4	4	4[a]	3	2	3
How often have you posted messages to online discussion forums for academic purposes?	2.8	3.1	2.4	3	3	2.5	3	3	1	3	2	3
How often have you used an online library catalogue system?	3.5	3.8	3.3	3	3	3	3	3	3	4	2	4
How often have you read descriptions of academic research projects?	3.0	3.3	2.8	3	3	3	3	3	2[a]	4	4	4
How often have you accessed websites of academic journals?	3.4	3.6	3.2	3	3	3	3	3	3[a]	4	2	4
How often have you accessed online discussion forums that are not associated with a course webpage?	2.8	2.7	2.8	2.5	3	2	2	3	2	4	3	4

Note. Questions were Likert-scaled, with 1 = very poor and 5 = very good for questions beginning with "How would you rate..." and 1 = never and 5 = very frequently for questions beginning with "How often have you..." ND = students without disabilities; LD = students with learning disabilities.

[a]Multiple modes exist. The smallest value is shown.

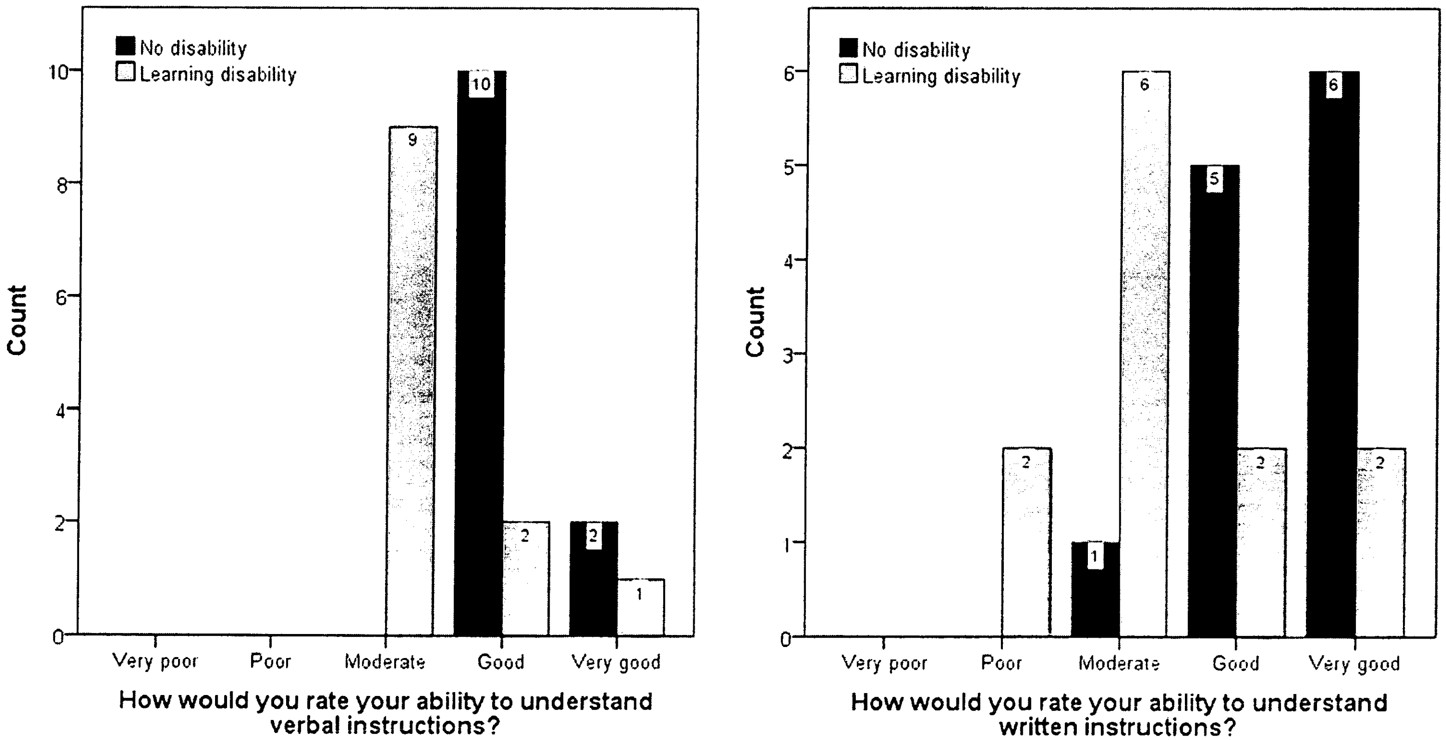


Figure 3. Frequency counts of responses of student participants with and without disabilities regarding their self-reported abilities to understand instructions.

4.2 Research Objective 1

Research Objective 1 was to determine the extent to which objective measures of the accessibility of e-learning technologies are able to predict the subjective accessibility experience of students. As previously summarized in section *3.6.4 Data analysis relevant to Research Objective 1*, data analysis relevant to this research objective included determining the precision of the data from objective accessibility testing, and considering the value of additional data obtained about module accessibility and general e-learning accessibility from student participants during subjective testing sessions.

4.2.1 Precision of objective data. Refer to Tables 6 and 7 for a complete list of all predicted and actual critical incidents for both modules, as determined by objective and subjective accessibility testing. Based on the results of objective accessibility testing, none of the webpages included in the modules would be deemed WCAG 2.0 Level AA compliant. Objective accessibility testing was effective at identifying potential accessibility problems related to accessing content, most commonly as related to the use of assistive technologies. For example, several instances of absent ALT text for images were highlighted across various webpages included in both modules. Potential critical incidents related to accessing an understanding of content were also frequently related to the use of assistive technologies. For example, incorrect use of header designations was identified on several module webpages, which could result in confusing output from text-to-speech software. Absent ALT text for hyperlinks is another example of how absent information could result in confusion for screen reader users.

Table 6.

Predicted and Observed Critical Incidents for the Credible Resources Module

Component	Predicted Critical Incident(s)	Observed Critical Incident(s) (Frequency)
Course homepage	Main Moodle logo lacks ALT text[a]	Unclear how to proceed from the course homepage (4)
Module homepage	Icon within the navigation menu lacks informative ALT text	Unclear how to proceed from the module homepage (3)
	Hyperlinks lack ALT text	YouTube video loading problems (1)
	Insufficient text-to-background color contrast for some text on the page	Unclear how to enlarge the YouTube video (1)
	Images and tables within PowerPoint presentation lack ALT text	Confusing PowerPoint slideshow controls (1)
	Duplicate PowerPoint slide titles	Too much information in the PowerPoint presentation (1)
		Information in the PowerPoint presentation is difficult to understand (1)
External website (academic)	Missing Search button label and form labels for AT use	Unable to locate information on the page (2)
	Search options are not grouped together for AT use	Unclear hyperlink text (1)
	Incorrect use of header designations for text on the page for AT use	

External website (non-academic)	Main page logo and banner ad lack ALT text	No observed critical incidents
	Many locations with insufficient text-to-background color contrast	
	Incorrect use of header designations for text on the page for AT use	
Moodle forum homepage	Unlabelled control area for forum display options for AT use	"Reply" button is not readily visible (1)
		Unclear how to post a comment (2)
		Unclear forum instructions (2)
Moodle forum reply page	Empty label for "Post to Forum" button for AT use	Unclear functionality of html editor buttons (1)
		Unable to paste text into a message (1)
		Unclear to how hyperlink text within a message (1)

Note. AT = assistive technology.
[a]This accessibility problem was predicted for all Moodle pages in both modules. For brevity, this problem is reported only once in the table.

Table 7.

Predicted and Observed Critical Incidents for the Scholarly Resources Module

Component	Predicted Critical Incident(s)	Observed Critical Incident(s) (Frequency)
Module homepage	Icon within the navigation menu lacks informative ALT text	Unclear how to proceed from the module homepage (1)
	Hyperlinks lack ALT text	Unclear module instructions (14)
	Insufficient text-to-background color contrast for some text on the page	Visually overwhelming module homepage (1)
	Images and tables within PowerPoint presentation lack ALT text	Unresponsive YouTube video "Play" button (1)
	Duplicate PowerPoint slide titles	Text in YouTube video is too small (2)
	Improper reading order on PowerPoint slide for AT use	Unclear how to enlarge the YouTube video (4)
		Once enlarged, the YouTube video freezes (2)
External websites (academic)	Library homepage: Main logo and page graphic lack ALT text; page language is not specified for AT use	Library homepage: No observed critical incidents
	Catalogue search results: Multiple checkbox buttons are not grouped and some checkbox labels for AT use are missing; incorrect use of header designations for text on the page for AT use; page language not correctly specified for AT use	Catalogue search results: Too many links on the page (4); unclear hyperlink purpose (6); unclear that different types of resources are described on the page (2); unclear what the information for individual resources represents (3)
	Full record for a specific catalogue search result: Main page logo and other image elements lack ALT text; text input boxes lack labels for AT use; page language is not specified for AT use	Full record for a specific catalogue search result: Unclear what the page is presenting (6); unclear which link to click on to access the resource (3)

External website (non-academic)	Main logo lacks ALT text	Unclear what the information on the page represents (4)
	Feature on the page is missing a label for AT use	
	Insufficient text-to-background color contrast for a page feature	
	Incorrect use of header designations for text on the page for AT use	

Note. Refer to Table 6 for an overview of predicted and observed critical incidents for the course homepage and Moodle forum pages, which were the same for both modules in the study.

Absent labels for buttons and form elements on webpages were also highlighted as potential critical incidents, which may affect the ability of students using screen readers to perform tasks (e.g., search for information) on webpages included in the modules.

Subjects who participated in this study do not normally use assistive technologies when working with webpages (as revealed by recruitment interviews) and did not use assistive technologies during this study. Therefore the majority of the potential critical incidents highlighted by objective testing were not applicable to them. As a result, even if the potential accessibility problems identified by objective testing methods were corrected/removed such that each page became WCAG 2.0 Level AA compliant, the accessibility of the webpages would not be expected to increase substantially from the perspective of the study participants. The sets of predicted and observed critical incidents were disjointed, and therefore the number of false negatives for each webpage was found to be equivalent to the number of unique critical incidents observed. For the Scholarly Resources module there were 14 unique critical incidents observed, while for the Credible Resources module there were 15 unique critical incidents observed.

The most common type of critical incidents identified from subjective accessibility testing were those related to accessing an understanding of content. For example, students did not always understand the purpose of hyperlinks or instructions included within the modules. While critical incidents related to accessing content (e.g., encountering a YouTube video with text that is too small to read) or performing tasks (e.g., difficulty locating a "Reply" button on a discussion forum) were frequently overcome by trial-and-error, this was not the case with critical incidents related to understanding. Moreover, critical incidents related to understanding often had a large impact on the ability of the students to succeed. Indeed, six students did not successfully complete the Scholarly Resources module as a result of confusing instructions included as part of the module homepage and/or confusing library catalogue search results.

4.2.2 Additional subjective data about module accessibility. Additional data regarding module accessibility was obtained from OpenVULab screencasts, video camera recordings, post-session questionnaires, and exit interviews.

4.2.2.1 OpenVULab screencasts and video camera recordings. The creation of screencasts using OpenVULab and the use of a video camera during subjective accessibility testing allowed for the collection of additional data, namely sentiment displays (verbal or non-verbal expressions of frustration or pleasure) and student-computer interaction. Nearly one-half of the participants in the study ($N = 11$) verbally expressed frustration while completing the Scholarly Resources module, while approximately one-fifth of the participants ($N = 5$) verbally expressed frustration while completing the Credible Resources module. For example, several students expressed frustration while attempting to locate online access to a journal webpage with comments including "Oh God, now I'm lost," "OK, why is this not working," and "Ugh!" When she did access the journal homepage, one student commented "This is going to take forever" upon observing the large amount of text presented on the page.

Verbal frustration data did not on its own lead to the identification of critical incidents, but it did (along with other think aloud statements) help to further highlight and/or confirm when a student was encountering difficulty and why. For example, the navigation pathway that students took when confused about how to access the journal homepage suggested that students were unsure of how to use the library catalogue search results, and additional think aloud data (including verbal expressions of frustration) helped at times to identify why the search results were confusing (e.g., uninformative hyperlink text). However, in most instances where verbal frustration was expressed (*N* = 13) this took place only once during completion of a module, so other cues (e.g., indirect navigation pathways and clicking on links that would not take the student to the webpage that they indicated via think aloud verbalization they were looking for) were often relied upon when identifying critical incidents.

Non-verbal expressions of frustration captured by the video camera were infrequent and not particularly helpful. In total, 22 video camera recordings were reviewed (one per student, omitting one student who declined to be video-taped and one instance where the researcher forgot to turn the camera on; 5 hours and 6 minutes of total video) and only four non-verbal expressions of frustration were noted in total. Three of these four instances coincided with observations from the OpenVULab screencasts (e.g., verbal frustration) and the fourth instance – a student tapping her fingers on her arm while waiting for a response from the computer – suggested frustration with a slow loading Moodle page. The fact that the page was loading slowly was also evident from the OpenVULab screencast.

Of the 48 OpenVULab screencasts (two screencasts per student, one per each module completed) and 22 video camera recordings reviewed, only two sentiment displays related to expressions of pleasure were noted. Two students made one comment each about attributes of a YouTube video and PowerPoint presentation that they liked, including when the video display zoomed in to better show information and the fact that the media were embedded on the page, making them easy to access. The low incidences of verbal expressions of pleasure may be related to a decreased likelihood of students expressing pleasure (i.e., perhaps students are more accustomed to verbally expressing frustration), the think aloud practice session in which the researcher modeled only verbal expressions of frustration, or a combination of both. While low in frequency, this information obtained from the OpenVULab screencasts may be deemed highly valuable as learning from students about features of the modules that enhanced accessibility may be useful when considering how to best address/prevent accessibility problems. Objective accessibility testing is typically designed to report problems rather than positive attributes of technologies, and the only other source of positive feedback about module accessibility was obtained from moderated post-session interviews.

Video camera recordings also allowed for detection of student-computer interaction that lent insight into module accessibility. The relevant interactions observed in this study were students leaning forward (i.e., closer to the computer monitor) to read text. This was observed taking place in seven unique instances (with five different webpages, a Powerpoint file, and an embedded YouTube video). This information was useful as, with the exception of the YouTube video and one comment about the text on the module homepages, students did not verbally express frustration via the think aloud protocol with small font size on webpages. This may be related to increased familiarity with user customization of YouTube videos compared to webpages as the students indicated that they knew that there should be a way to enlarge the video, and/or decreased resolution of the text in the YouTube video (i.e., rending the small text in the video more problematic than the small text on the webpages). It is not known whether students knew that they could manipulate the web browser to enlarge the font size on webpages.

4.2.2.2 Post-session questionnaires. Refer to Table 8 and Table 9 for summaries of the post-test questionnaire results for each module, including the mean, median, mode, and range for the responses obtained for each question. While the tabulation of predicted and actual critical incidents (Table 6 and Table 7) allowed for a determination of the amount and nature of unique accessibility challenges encountered, the post-test questionnaire results offer further insight into the impact of those challenges. Objective accessibility testing predicted 19 and 14 critical incidents for the Scholarly and Credible Resources modules, respectively, whereas the actual number of unique critical incidents observed from subjective accessibility testing were 14 and 15, respectively. These data may suggest that the two modules may be considered somewhat similar in terms of an overall accessibility rating. However, student rating of the ease of completion of the modules suggested otherwise.

Table 8.

Participant Feedback from Questionnaires Administered After Completion of the Scholarly Resources Module

Question	Mean	Median	Mode	Range
Overall, how easy was it to work within Moodle while completing this module?	3.7	4	4	3
How easy was it to read the instructions included within this module?	3.7	4	4	4
How easy was it to understand the instructions included within this module?	3.6	4	4	3
How easy was it to access the PowerPoint presentation within this module?	4.4	5	5	3
How easy was it to understand the PowerPoint presentation within this module?	4.0	4	4	3
How easy was it to access the video included within this module?	4.3	4.5	5	3
How easy was it to understand the video included within this module?	3.8	4	4	4
How easy was the question included at the end of this module?	3.8	4	4	4
How easy was it to post your answer to the question on the discussion forum in Moodle?	4.3	4	4	2

How easy was it to work with websites outside of Moodle while completing this module?	3.4	3	3	3
Overall, how easy was it to complete this module?	3.5	4	4	3
Overall, how easy was it for you to participate in this testing session?	3.9	4	4	3
Overall, how would you rate your comfort level with participating in this session?	3.8	4	3	3

Note. Survey questions were Likert-scaled, with 1 = very difficult and 5 = very easy for questions beginning with "How easy was it…" and 1 = very low and 5 = very high for the question beginning with "Overall, how would you rate your comfort level…"

Table 9.

Participant Feedback from Questionnaires Administered After Completion of the Credible Resources Module

Question	Mean	Median	Mode	Range
Overall, how easy was it to work within Moodle while completing this module?	4.2	4	4	3
How easy was it to read the instructions included within this module?	3.8	4	4	4
How easy was it to understand the instructions included within this module?	4.1	4	4	3
How easy was it to access the PowerPoint presentation within this module?	4.5	5	5	2
How easy was it to understand the PowerPoint presentation within this module?	4.0	4	4	4
How easy was it to access the video included within this module?	4.2	4.5	5	3
How easy was it to understand the video included within this module?	4.2	4	4	3
How easy was the question included at the end of this module?	4.2	4	4	2
How easy was it to post your answer to the question on the discussion forum in Moodle?	4.4	5	5	3
How easy was it to work with websites outside of Moodle while completing this module?	4.2	4	4	3

Overall, how easy was it to complete this module?	4.3	5	5	3
Overall, how easy was it for you to participate in this testing session?	4.1	5	5	4
Overall, how would you rate your comfort level with participating in this session?	3.8	4	4	3

Note. Survey questions were Likert-scaled, with 1 = very difficult and 5 = very easy for questions beginning with "How easy was it…" and 1 = very low and 5 = very high for the question beginning with "Overall, how would you rate your comfort level…"

Answers to the question "Overall, how easy was it to complete this module?" following completion of the Scholarly Resources module had a mean, median, and mode of 3.5, 4, and 4, respectively. When the same question was answered following completion of the Credible Resources module, the mean, median, and mode were found to be 4.3, 4.5, and 5, respectively. To determine whether these responses were significantly different across the two modules, the Wilcoxon signed ranks test was conducted to compare the median responses. This non-parametric test was chosen as an alternative to the paired *t* test as the data was found not to exhibit normality. The analysis revealed that there was a significant difference in the students' perceptions of ease of the modules ($Z = -3.043$; $p = 0.002$) with the Scholarly Resources module deemed more challenging. For a visual depiction of how this data differed across the two modules, refer to Figure 4 which presents the data in the form of clustered bar charts. Refer also to section *4.3.1.2 Critical incident counts* where comparison of critical incident counts across the two modules is discussed, and where this assertion is further supported.

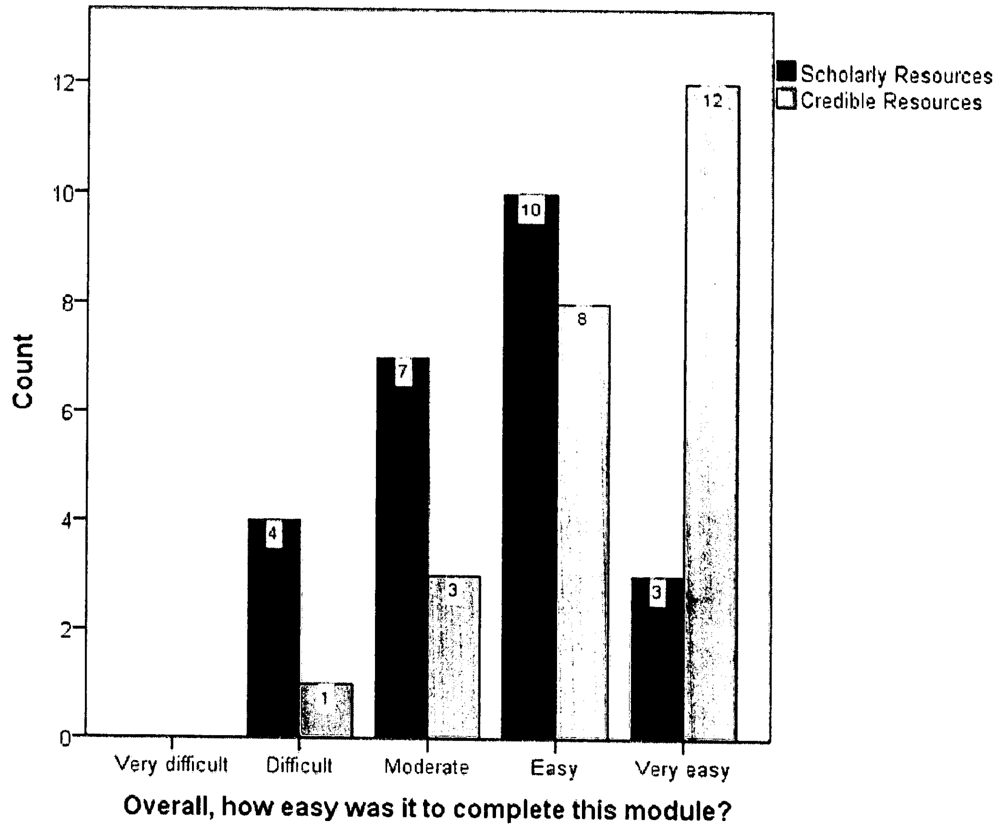


Figure 4. Frequency counts of student perceptions of the relative ease of completion of the Scholarly and Credible Resources modules.

Further examination of responses for the 10 questions related to specific module components provided additional insight into the students' perceptions of the accessibility of each module. From the post-test questionnaire results for the Scholarly Resources module (Table 8), students rated ease of working with websites outside of Moodle lower than all other aspects of module completion. Second to this was understanding the module instructions. These ratings coincide with observed critical incident data from this module (Table 7), in which it was observed that students misunderstood module instructions and were unsure of how to work with library catalogue search results. These post-test questionnaire data help to confirm the critical incident data because, as discussed in the previous section, students did not always express frustration when they encountered difficulty and it was necessary at times to make inferences about the nature of problems that were observed in the OpenVULab screencasts. On the other hand, the post-test questionnaire did not reveal that many students struggled with enlarging a YouTube video that contained small text, as was noted from the OpenVULab recordings. This discrepancy may be related to how the students interpreted the question, as students may have considered only the ease with which they played the video using the player controls and not the ease with which they could access the content presented in the video when responding to the question.

The post-test questionnaire results for the Credible Resources module revealed that most students found nearly all aspects of this module to be "easy" or "very easy." While there were critical incidents observed for all but one website included within this module, the questionnaire data coincides with the observations from the OpenVULab screencasts in that the challenges encountered did not prevent students from completing the module and were thus overcome by the students (all of whom successfully completed the module).

4.2.2.3 Exit interviews. In the exit interviews, students identified the most and least challenging aspects of each module. Refer to Table 10 and Table 11 for a complete list of the responses for the Scholarly and Credible Resources modules, respectively.

Feedback from students regarding challenging aspects of the Scholarly Resources module was aligned with observations made from the OpenVULab screencasts and/or the post-test questionnaires. The most common comments were that students found it challenging to understand how to work with library catalogue search results (i.e., websites external to Moodle), to view a YouTube video that was too small, and to understand module instructions. In addition to confirming the aspects of the module that most students found particularly challenging, the interviews also revealed additional aspects that some students found to be particularly problematic that weren't revealed by other sources of data. For example, it was only from the interviews that it was revealed that the terminology in the module was challenging for some and that it was difficult for some to stay engaged for the duration of the YouTube video.

Several students commented that they did not find any aspect of the Credible Resources module to be particularly challenging and that the only issue for them was that they were perhaps less engaged due to the lack of challenge. As students identified challenges less readily from this module, it is not clear how impactful the identified challenges were. It is therefore not surprising that there is not a strong connection between these data from the interviews and the post-test questionnaire in which students rated all aspects of the module quite favorably. However, the interview data for this module also provided insight into factors that could impact accessibility to some extent for some students that were not identified by other sources of data. For example, the inclusion of websites external to Moodle with a variety of layouts was rated as a challenge by some students.

Table 10.

Most and Least Challenging Aspects of the Scholarly Resources Module as Revealed by Exit Interviews

Most Challenging Aspects (Frequency of Response)	Least Challenging Aspects (Frequency of Response)
Understanding what to do with library search results (6)	Playing embedded media (7)
YouTube video was too small (4)	Simple, step-wise module layout (5)
Understanding module instructions (2)	Formulating an answer to the question in the module (3)
Unfamiliar journal website layout (1)	Using the Moodle discussion forum (2)
Unexpected/unfamiliar module layout (1)	Similar structure to the other module (2)
Applying learnings from module to locating a journal article (1)	Understanding the PowerPoint and YouTube video (1)
Too much information on the module homepage (1)	The ability to review information (1)
Reading all of the module instructions and understanding expectations (1)	Getting library catalogue search results (1)
Too much text in the module (1)	Reading module instructions (1)
Not enough instruction for using library catalogue (1)	Short, concise, and clear PowerPoint slides and YouTube video (1)
No access to help (1)	Familiarity of the library catalogue (1)
Using the PowerPoint slideshow controls (1)	Familiarity with the module topic (1)
Challenging terminology used within the module (1)	
Small font on the module homepage (1)	
YouTube video was too lengthy to stay engaged (1)	

Table 11.

Most and Least Challenging Aspects of the Credible Resources Module as Revealed by Exit Interviews

Most Challenging Aspects (Frequency of Response)	Least Challenging Aspects (Frequency of Response)
Staying engaged (4)	Playing embedded media (5)
Formulating an answer to the question in the module (3)	Applying learnings from the module to external websites (2)
Working with external websites with different layouts (3)	Similar structure to the other module (2)
A lot of information presented in the YouTube video (2)	Helpful and understandable YouTube video (2)
Uncertainty about expectations (2)	Short, concise, and clear slides and video (1)
Cannot enlarge the YouTube video (1)	Familiarity with the external websites (1)
Unsure how to play the slideshow (1)	Familiarity with YouTube videos (1)
Unexpected/unfamiliar module layout (1)	Using the Moodle discussion forum (1)
Too much information all at once (1)	Step-wise module layout (1)

Examination of the least challenging aspects reported for each module revealed several similarities across the two modules. Students frequently highlighted the fact that the media (PowerPoint slideshow and YouTube video) were embedded onto the module homepage as a favorable attribute. The simple step-wise layout of the module was also popular, as was the similarity in structure of the two modules. Familiarity with other components of each module was also highlighted as helpful. These data shed more light on factors that influenced the accessibility of the modules that were not identified by other sources of data which were perhaps more suited to identify accessibility problems vs. attributes that enhanced accessibility.

In addition to identifying most and least challenging aspects of each module, students made additional comments about attributes of technologies included in each module that influenced their perception of module accessibility. Several themes were identified: familiarity, simplicity, and engagement and learning style, and each are discussed below.

4.2.2.3.1 Familiarity. Familiarity emerged as a theme relevant to all aspects of the modules (Moodle webpages, PowerPoint slideshows and YouTube videos embedded within the module homepages, and external academic and non-academic webpages). In general, familiarity with technology included within the modules and the structure of the modules themselves contributed positively to students' perceptions of accessibility, while unfamiliar module structure or components were problematic for students.

Aliya's comment about the familiar structure of the second module that she completed (in comparison to the first module she completed) sums up sentiments expressed by several students:

> That it was the same format as the first one [was helpful]. So I didn't really pay too much attention to like how it was laid out because it seemed to be, like the instructions were the same, so I was a little bit more comfortable and a little more at ease in terms of like just um going through and finding everything.

These comments also highlighted the value of the counterbalancing of the study design in that half of the students completed the Scholarly Resources module first, while the other half completed the Credible Resources module first.

Other students indicated that their ability to complete specific aspects of the modules in Moodle was also affected by familiarity. For example, Ollie found that she was unsure what was required of her when completing the modules "'cause it's different. I'm not used to completing things like this." Mandy, on the other hand, indicated that responding to the culminating questions in the modules was pretty straightforward as she was "used to questions like that" and had posted to discussion forums before.

Students also described prior familiarity with PowerPoint and YouTube videos. Aliya stated that "I'm used to YouTube so um, it was easy for me to kind of maneuver through that and know how to use the settings at the bottom." At the same time, past experience with viewing YouTube videos did not always benefit the students as they expressed frustration at knowing that there should be a mechanism to enlarge the video but were unable to follow steps that were familiar to them. According to Carla, "usually there is an option that says full screen but maybe not when you embed it [a YouTube video] on the Moodle site." Morgan and Arla were aware of an alternate strategy that would have allowed them to enlarge the embedded video, but chose not to because of familiarity with new challenges that could arise as a result. For example, Arla stated:

> I could've maximized it [the YouTube video], but I didn't know – I know that sometimes if you maximize a YouTube video it freezes, so.... Yeah, because I know, like if you change it from like a small one and maximize it for YouTube,

like a proper size, or take it to the website, it sometimes freezes. Especially at York, I've seen it happen a lot. So I didn't want to do it.

From these examples, it is evident that familiarity with features of e-learning infrastructure can have an enabling or disabling impact on students, as this can make it easier for students to access learning materials or may cause them to avoid customizing their learning experience for fear of re-experiencing challenges encountered in other contexts.

The impact of familiarity and prior experience also emerged as a theme relevant to websites external to Moodle that students encountered. One such website was an online academic library catalogue that students were asked to work with. Brad found that familiarity with this type of technology helped him to complete the relevant module:

> Overall I'd say the tasks involved [in completing the Scholarly Resources module] weren't too challenging because I've already used the York catalogue so this is very similar to that. So I just kind of transferred whatever I do there to this.

On the other hand, instances where students encountered unfamiliar technologies or webpages increased the level of challenge for students. Mandy found working with the online library catalogue search results more challenging than Brad did, as she stated:

> So I went back to it, but it took me to this page, and I've never been to anything like this. So I clicked on this because it said "full text available." And then it took me to the website. And I was like, oh, it's a separate website. I didn't know that. It confused me.

Another point raised by students was that it can be challenging to work with multiple webpages that have different layouts because similar information may be placed in different locations on the websites.

4.2.2.3.2 Simplicity. Simplicity of e-learning infrastructure was found to have a positive impact on students' perceptions of accessibility. Many students commented positively on the simple step-wise layout of the module homepages. Consider Leanne's thoughts on this:

> I think like the way it's laid out – it's very clear. Like you have the PowerPoint presentation. Like you don't have to, you don't have to download it, open it on to your own computer and search it, like it's embedded right into the page which is great. Um and it's laid out clear like your Step A, Step B which is nice. And then it's the exact thing with the YouTube clip, like it's embedded right into the page. You don't have to, you know, click the link um or like sorry you don't have to type in the link to YouTube and search for it yourself.

Indeed, several students highlighted the value of having the media embedded on the page and the benefit of not needing to view multiple screens in order to access all of the module instructions. However, Penny found this same attribute to be a challenge for her as:

> I find that I get distracted very easily. And so, when all of the instructions are listed out at the same time I find it difficult and somewhat overwhelming at first, when I first look at it. Um, so I would probably prefer one-at-a-time to get the instructions maybe, this page, and then click the next page.

This contrasting view points to the subjective nature of accessibility, in that what is most accessible to one student may not be for another.

While simplicity of the module homepage structure was most often highlighted in a positive manner by the students, many expressed frustration that other aspects of module completion were more complex. For example, accessing an embedded YouTube video was convenient but the videos themselves were complex in that they included a lot of information. For example, Donna, Janna, and Julie found that this made it difficult for them to complete the Credible Resources module because they could not remember all of the tips highlighted in the video when subsequently attempting to assess credibility of an external website. However, it is unclear whether this same challenge would have been experienced in an authentic e-learning experience where students may take notes while watching the video that they could refer to afterwards while applying new knowledge from it. In this study, students were not prohibited from taking notes, though none did (nor did any students ask if they could take notes).

Complexity of external websites was also a challenge for participants, and several students discussed frustration with the online library catalogue search results. According to Mia, "Um, I guess the last part where you have to find the scholarly article [was most difficult] because there were so many search engines [appearing in the library search results]. I kind of got confused." Similarly, Connie commented that:

> There were too many links [on the catalogue record for the journal]. So it's like, you don't know what to click and then once it gave you one link, you're

> wondering if you clicked the wrong link, or.... Or is it somewhere on the page and you can't find the new page.

Indeed, the online library catalogue webpages were the sources of many critical incidents when the OpenVULab screencasts were reviewed.

4.2.2.3.3 Engagement and learning style. Several students offered comments related to the theme of engagement when reflecting on their experience completing the modules. For example, David had difficulty staying engaged and thus accessing information from the PowerPoint presentation included within the Scholarly Resources module:

> Yeah, I was like, uh, blah blah blah [when reading the table in PowerPoint]. (laughs).... like, let's just get on with it.... Um, so I just kind of skimmed through. I was like, OK, peer review, they're expert written, they have a scholarly tone, and they cite their references. I was like, OK. So if this said anything important, I probably would have missed it.

For Mandy, on the other hand, it was the Credible Resources module that she felt less engaged with. This is because she felt that she was already familiar with much of the content presented.

The length and quality of e-learning materials was also highlighted as a variable that can influence engagement of the students. Janna compared the videos in the modules to screencasts that her professor prepared for an online course that she was currently taking:

> Um, this video specifically is um is very good. It's loud and clear. It's not monotonic like the one that I'm taking now. The professor makes us fall asleep and we can't sit there and every lecture, there's two lectures a week and they're each about an hour or more and uh, it just gets very very tiring.

While Janna liked the YouTube videos included in the modules in this study, Leanne found that they were too lengthy for her to remain engaged:

> I was not really a fan of the [Scholarly Resources] video and I found it was just too lengthy and I don't think I would necessarily sit through three minutes or three or four minutes of listening to what that was about.

These contrasting views about the same e-learning infrastructure also points to the value of providing students with multiple means of representation of course content so that all students may have access to materials that best suit their learning preferences and needs.

Indeed, many of the study participants expressed a high degree of self-awareness about their learning preferences and needs, and the inclusion of learning materials that best suited them as learners was highlighted as valuable. To this end, Karen, Penny, and Susan described themselves as "visual learners" and identified the visual components included within the PowerPoint presentation and YouTube videos as aspects of the e-learning infrastructure that they found particularly helpful. Penny would have also liked to have had access to the module instructions in audio format, as she commented that:

> I think that sometimes instructions, like these instructions if they were given out loud I may have been a little quicker to understand them rather than stopping to

read them and reading them again and again I guess? I guess I'd also appreciate the audio if it was there.

In contrast, Morgan stated "I'm a very literal learner, so it's more easy for me to follow instructions when they're written down than verbally or visually." Again, these examples suggest that a variety of means of representation of content within the modules would therefore help to maximize accessibility for the participants in this study.

4.2.3 Subjective data about e-learning accessibility. In their discussion of specific attributes of the e-learning infrastructure encountered in this study, several themes also arose from review of exit interview transcripts that are relevant to the enabling or disabling impact of e-learning technology in a more general, broader sense. The themes of engagement, learner control, tolerance, and social presence emerged, and are each discussed below.

4.2.3.1 Engagement and online learning. Students expressed a variety of views regarding their engagement with specific components of the e-learning infrastructure encountered in this study. In addition, engagement in online learning environments was discussed in a broader sense. Some students speculated that engagement with the learning process may increase in an online learning environment, while others speculated to the contrary.

Jeremy and Mandy both suggested that the possibility of interactive online learning environments may enhance their engagement compared to face-to-face learning. According to Jeremy, "if we were in a classroom you'd be watching the professor doing it. This [online learning] is probably more interactive." Mandy's thoughts on interactivity are from the perspective of a student with a learning disability:

> If they have a learning disability and this [Moodle] website was used for them, for example, a lot of people when they're being taught they don't pay attention as much and they need the computer for interacting. I don't know, it's, it gives them stimulation usually. I find. Umm, so maybe being able to read it could give them the opportunity to pay more attention.

Mandy went on to say "It's just, looking at her [my professor in a face-to-face class], it just distracts me. Just looking at them and not doing anything distracts me. I don't know. I have to take notes or something." Mandy's comments point to a relationship between engagement and distraction, which other students also highlighted. Like Mandy, Donna also finds that face-to-face classes can include attributes that are more distracting for her compared to online learning environments. For Donna, it is the presence of other students in the classroom that she finds distracting and which can take detract from her ability to focus on her work.

However, in contrast to the point of view of Mandy and Donna, Mia finds that she is more apt to become distracted when working online. For Mia, being in the presence of a "live" professor has a positive impact on her engagement with the learning material. Allison expressed similar feelings when she said "Because when it's face-to-face, well, if you're showing me the information and speaking to me, I think that would have been a little more easier [sic] because um, I would be more in tuned to listen to you." Similarly, Janna speculated that "in class she's [the professor] there, she has to be more uh, I don't know, interactive, more passionate, more, you know..." The large number of comments related to the theme of engagement in online vs. face-to-face learning environments points to a high degree of importance placed on this variable by students. At the same time, the variety of responses about factors that influence engagement highlight the need to incorporate multiple means of engagement into learning environments to meet the needs of diverse populations of learners.

4.2.3.2 Learner control and online learning. Students indicated that an enhanced sense of learner control can be an enabling attribute of online learning. Students expressed that they appreciate the ability to take enhanced control over how and when they engage with the learning material (i.e., to be more autonomous learners).

For Allison, the independence associated with online learning can be empowering. She stated:

> Yeah, it was more like the independent piece [that I liked about completing the modules]. Like, um, it kind of just… I don't know. I kind of felt like a big person in a sense? Like kind of important like, you know, I'm just doing my thing.

For many students in this study, "just doing my thing" includes the ability to work at their own pace, including taking a break when needed, and the ability to review learning materials that are available online as needed.

Choosing when to engage with learning materials was highlighted as a benefit of online learning compared to face-to-face classes which may be scheduled at times that are not optimal for all learners. For example, Penny discussed this when she said:

> I mean in terms of face-to-face classes you may not be ready mentally for that class. Whereas if you were to have an online class you sort of, you choose when you're ready to start your work or when you're ready to listen to the – I've never had one before but, I mean, to listen to the lecture or whatever it is that's posted.

Janna does have experience with online learning, and confirmed that this attribute is useful to her as she noted that:

> Yeah, well, there is the advantage of well, OK, I can wake up at 3 o'clock in the morning and I can do the lecture or whatever and I don't have to um, you know, think about when to go to class. I don't have to be anywhere.

Similarly, David takes almost exclusively online courses and finds the ability to work in smaller chunks of time helpful for him because he finds it difficult to sit still during lengthy face-to-face classes and is frustrated when he cannot "rewind" what is said in the face-to-face environment.

In contrast, Morgan expressed concern that online learning has the potential to be more controlling than face-to-face. When reflecting on her experience in completing the modules in this study, she commented:

> Yeah, I can definitely see people not wanting to go through instructions.... It's quicker when someone tells you to do it. Yeah, and you're like, OK fine I got it and you just go do it instead of having to go OK, step by step. Because when you have steps laid out for you, you get the feeling that you have to read all of them. Just follow exactly what they're doing.

Morgan's comments suggest that the structure of an online learning environment may have an impact on the degree of freedom that students may feel with respect to how to engage with the course, and that allowing students to explore multiple methods of exploring and expressing their learning is valuable.

4.2.3.3 Tolerance and online learning. Several students indicated that online learning environments may be more tolerant of individual learner needs compared to face-to-face environments. Students in this study expressed concern that their individual needs may be disruptive in a face-to-face class and that they may at times feel more comfortable to seek answers to questions when working online. For example, Julie noted that:

> Sometimes in class you might miss something. In a lecture hall of 100 students or more, you feel apprehensive of raising your hand and asking the professor to repeat it. So in regards to doing an online class, again it's me by myself. If I forget something while I'm doing a task, I can go back and say, OK what did I forget?

Similarly, David and Allison expressed apprehension at interrupting the flow of a face-to-face class to ask questions. According to David:

> I tend to be impatient so whenever there is a question I just start opening up windows and like Googling it and stuff. Whereas if I'm in a class, there's a lot of people there, and I feel like I'm slowing it down and I don't want to ask anything.

David went on to say "When he [the professor] does give you an answer, you can't say, I still don't get it. You gotta keep this rolling, right?" Allison commented on her experience determining how to enlarge the YouTube video in one of the modules in this study. She was able to determine how to enlarge the video by trial-and-error, though indicated that she would not have felt comfortable interrupting a face-to-face class to ask for assistance with that.

Mandy, on the other hand, speculated that online learning environments that rely heavily on text-based communication could be less tolerant of learner differences. For example, she mentioned that "obviously, you can't hear the person's tone of voice online. There could be problems with that too – misunderstandings." Her comment suggests that some learners may have more difficulty than others in expressing themselves in writing and/or understanding/interpreting communication by others that is presented in written form.

4.2.3.4 Social presence and online learning. Many students in this study expressed concern about a possible reduction in social presence in online learning environments compared to face-to-face environments that could negatively impact their engagement and success in a course. For example, some students felt that it may be easier to develop rapport with the teacher and to understand the teacher's expectations in a face-to-face setting. According to Eva, "you get to develop a real rapport and you get a better sense of who the prof is and what they really want [in a face-to-face setting]." Similarly, Carla and Peter indicated that some aspects of the learning process may work better for them in a face-to-face setting. Carla described challenges associated with engaging in and following a discussion with peers using an online discussion forum when she said:

> Like 'cause on the forum you have to wait, right? For the person to see it again. And then it's, let's say there are 30 students in the class. It's hard to um to follow

> up on what someone said because it doesn't – I think it doesn't go on – even if you click on reply at any spot, it [your message] goes at the end [of the conversation].

For Peter, it was interactions with the teacher that he highlighted as potentially being less effective online when he said "the positive things about it being in person is that if you have any questions that can't be answered online. The professor will understand you better and it's just more engaging when you are in person." Mia also noted that she feels that face-to-face discussion feels more personal compared to online discussion. Jeremy, however, indicated that "as long as the, whoever is running it – the professor – has adequate contact with the students I wouldn't feel a need for face-to-face."

Students also noted that the learning context may influence the suitability and the importance of different forms of social interaction. According to Jessica:

> I think it, like if it was social science or something like that it would be better to work with people and see their reactions. Because, for example, if you want to be a lawyer or something like that, there will be times when you have to face the person, right? And it's not the same to write a memo saying "You are guilty of..." (laughs)

Jessica's observation suggests that it may be helpful to consider which forms of communication may be most suited to particular learning objectives.

4.2.4 Summary of Research Objective 1 findings. Overall, objective accessibility testing was not effective in predicting the subjective experiences of the student participants in this study. Most of the critical incidents predicted by objective methods of accessibility evaluation were not applicable to the experiences of these participants, as they were related to the use of assistive software, which was not used in this study. Moreover, students encountered many challenges with accessing an understanding of e-learning infrastructure encountered in the modules while objective methods were less effective in predicting accessibility problems related to understanding (i.e., there were many false negatives related to understanding). Additionally, the challenges in understanding module content were most likely to have a large impact on the students' ability to complete the module, underscoring the importance of employing methods of accessibility evaluation that may identify such challenges.

In addition to an inventory of unique observed critical incidents, subjective accessibility evaluation yielded a substantial amount of additional data that further supported the identification of critical incidents, shed more light on the nature and impact of critical incidents, and aided in the identification of additional attributes of the e-learning infrastructure that impacted accessibility from the students' perspective. Post-session questionnaires highlighted the contrasting overall perception of accessibility of the two modules, which was not as readily evident from the results of objective evaluation. Not only was it evident from the post-test questionnaires that the Scholarly Resources module was deemed more challenging than the Credible Resources module, it was also possible to identify which aspects of the module students found to be most challenging. These data, along with think aloud verbalizations including verbal expressions of frustration, helped to better determine the nature and impact of the critical incidents that were observed in the OpenVULab screencasts.

Exit interviews provided additional information not only about the accessibility of specific e-learning infrastructure included in this study, but also about potential enabling and disabling attributes of online learning. Exit interviews were the primary source of feedback regarding attributes of the mock online course that had a positive impact on their ability to complete the modules. Students highly valued familiarity and simplicity within the modules, and highlighted the importance of being engaged with learning materials and working with materials that suit their preferred learning needs and style. Students also noted that being engaged in the learning process, feeling a sense of learner control, tolerance for error and individual exploration/expression, and social presence are attributes that would enhance their e-learning experience in a variety of contexts, not just as related to this study.

Finally, throughout the reporting of the results for this research objective, it is evident that accessibility is a highly individualized attribute. While there were common trends observed regarding accessible and inaccessible attributes of the e-learning infrastructure, there were also frequently contrasting views from the students about enabling and disabling features, and the number of critical incidents (if any) that individual students encountered when completing the same modules also varied. This observation supports the notion of imparting flexibility and variety into e-learning infrastructure in order to enhance accessibility for diverse learner populations.

4.3 Research Objective 2

Research Objective 2 was to determine whether data obtained from moderated and unmoderated e-learning accessibility testing are different and, if so, how and why they differ. As previously summarized in Figure 2, data analysis relevant to this research objective included comparison of the data obtained from review of OpenVULab screencasts (counts and efficiency of module completion), comparison of student perceptions of accessibility and preferences participating in accessibility testing across the two session formats, and a consideration of the value of additional data obtained only from moderated sessions.

4.3.1 Data from moderated vs. unmoderated sessions. Statistical analyses using SPSS were conducted to compare the efficiency of module completion and counts of critical incidents and verbal frustration across the two methods of subjective accessibility testing.

4.3.1.1 Efficiency of module completion. Efficiency of module completion was determined as the time (minutes) that it took each student to complete a module, and was tabulated following review of OpenVULab screencasts. Efficiency data points were collected for this experimental condition only in instances in which students successfully completed all aspects of a given module. Since six students did not successfully complete the Scholarly Resources module, only 18 efficiency scores were included for that module. All students successfully completed the Credible Resources module. However, it was deemed necessary to intervene to re-direct Aliya during completion of this module and therefore the time that was taken by Aliya to complete the module may have been extended as a result. This efficiency data point was omitted and therefore 23 efficiency scores were included in this analysis for completion of the Credible Resources module. Figure 5 presents a histogram of the efficiency data.

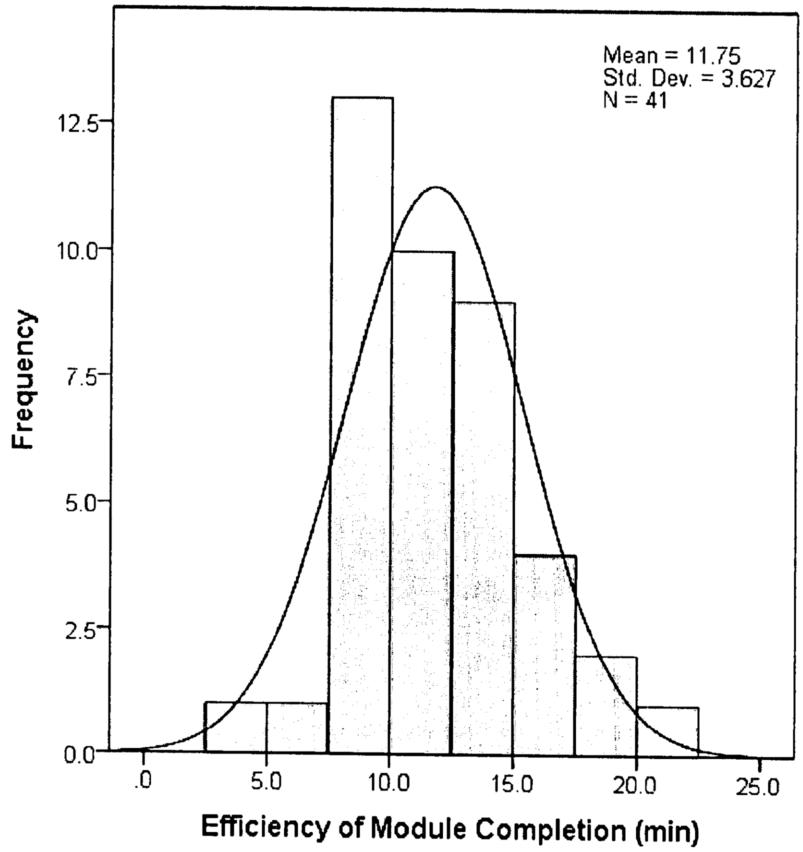


Figure 5. Histogram illustrating efficiency of module completion with a normal curve overlay.

The histogram in Figure 5 illustrates that the data for efficiency of module completion appear to be normally distributed. This was confirmed by the Shapiro-Wilk test for normality ($W = 0.968$; $p = 0.288$) and visual inspection of a Q-Q plot generated from the efficiency data. Standardized residuals were also examined for normality in a similar fashion (refer to Figure 6) and were also deemed to fit a normal distribution according to the Shapiro-Wilk test for normality ($W = 0.968$; $p = 0.288$) and inspection of a Q-Q plot.

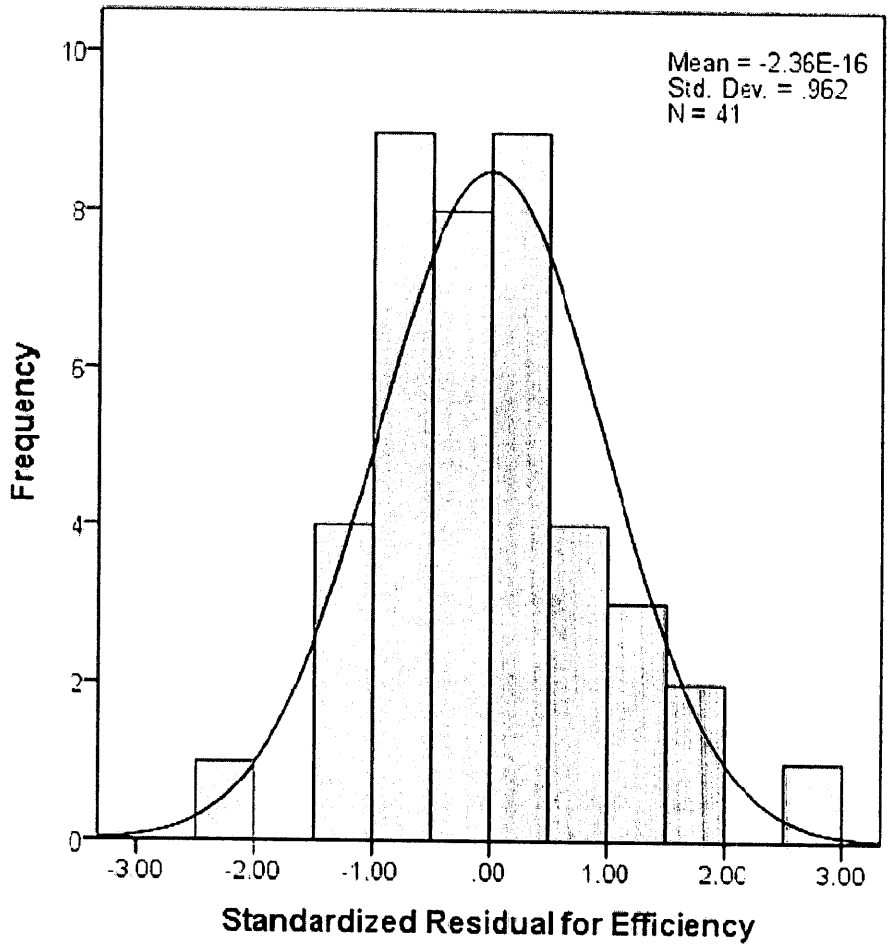


Figure 6. Histogram illustrating standardized residuals for the efficiency of module completion data with a normal curve overlay.

Levene's test for equality of variances was conducted and determined that the variances amongst the different testing groups were found not to differ significantly ($F = 0.207$; $p = 0.891$). With the assumptions of normality of data and homogeneity of variances met, two-way ANOVA analysis was carried out with format of testing session (moderated or unmoderated) and module (Scholarly or Credible Resources) as independent variables, and efficiency (minutes) as the dependent variable.

The two-way ANOVA analysis revealed that there was no significant effect on the efficiency of module completion attributed to the format of testing method ($F = 0.058$; $p = 0.810$) or the module completed ($F = 1.575$; $p = 0.217$), nor was there a significant interactive effect of the format and module on efficiency ($F = 2.666$; $p = 0.111$). To determine whether the presence or absence of self-reported disability may have had a significant impact on the efficiency data, the analysis was also repeated to examine for main effects or interaction with disability. However, no significant effect was found (with p values remaining well above 0.05), indicating that, in this e-learning context, the presence or absence of self-reported disability status did not significantly affect the time required for students to complete the modules.

4.3.1.2 Critical incident counts. Instances in which students encountered difficulty accessing module content, accessing an understanding of module content, or performing a task were recorded as critical incidents upon review of OpenVULab screencasts. The total number of critical incidents encountered by each student for each module was recorded, and Figure 7 presents a histogram of the critical incident count data.

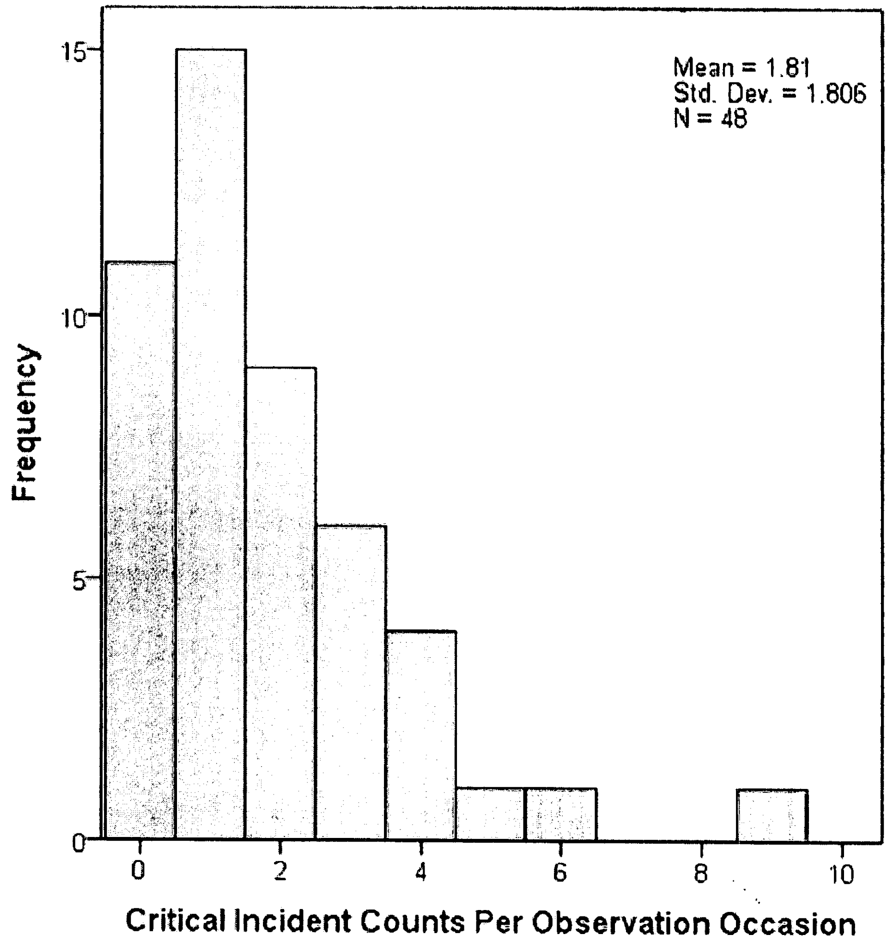


Figure 7. Histogram illustrating critical incident counts recorded during module completion.

From Figure 7 it is evident that these count data do not fit a normal distribution. The data do appear to resemble a Poisson distribution, and the mean and variance were compared to determine if they are approximately equal, a characteristic of this distribution (Upton & Cook, 2008a). The mean was found to be 1.81 and the variance was found to be 3.262. To further explore this, a scatterplot of the observed count frequency with an overlay of the expected count frequency for a Poisson distribution (predicted by SPSS from the mean of the data) was prepared. Figure 8 illustrates that the data fits a Poisson distribution fairly well.

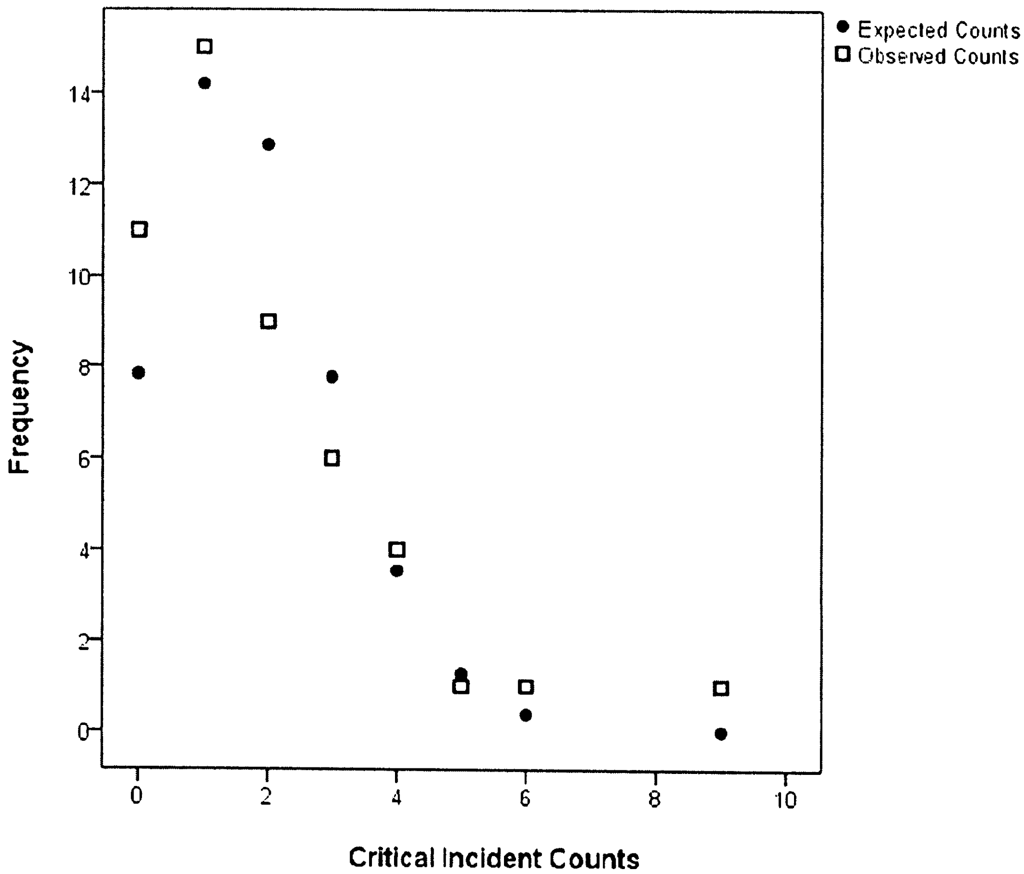


Figure 8. Scatterplot illustrating observed critical incident count frequency with an overlay of the expected count frequency for a Poisson distribution.

Based on the reasonable fit of the data to a Poisson distribution, Poisson regression analysis was conducted with format of testing session (moderated or unmoderated) and module (Scholarly or Credible Resources) as independent variables and critical incident counts as the dependent variable. This analysis revealed that the module had a statistically significant impact on the critical incident counts (Wald chi-square = 9.928; $p = 0.002$) with higher counts observed during completion of the Scholarly Resources module. This finding is aligned with the increased perception of difficulty of this module as reported by students in the post-test questionnaire (refer to section *4.2.2 Post-session questionnaires* and Tables 8 and 9), and the observation from OpenVULab screencasts that 25% of the participants ($N = 6$) did not successfully complete this module. The regression analysis also revealed that the format of the testing session did not significantly impact the count data (Wald chi-square = 0.134; $p = 0.714$) nor was there an interaction between format and module (Wald chi-square = 1.002; $p = 0.317$). In other words, there were higher counts of critical incidents from completion of the Scholarly Resources module regardless of whether the module was completed during a moderated or an unmoderated session. To determine whether the presence or absence of self-reported disability may have had a significant impact on the critical incident count data, the analysis was also repeated to examine for main effects or interaction with disability. However, no significant effect was found, indicating that, in this e-learning context, the presence or absence of self-reported disability status did not significantly affect the number of critical incidents observed.

Given that the data were found to be overdispersed for a typical Poisson distribution (with the variance exceeding the mean) and that the data may therefore fit a negative binomial distribution (Cameron & Trivedi, 1998), negative binomial regression analysis was also conducted. To determine which analysis (Poisson regression or negative binomial regression) was most appropriate, the Consistent AIC (CAIC) values obtained from each regression were compared, as a lower CAIC indicates a better fitting model (Hox, 2010). This comparison suggested that the Poisson regression may have been more appropriate for the data as the CAIC value of 168.985 was lower than that reported from the negative binomial regression (CAIC = 175.623). However, the observed trends with respect to the influence of format of testing session and module completed on critical incident counts were found to be the same from both regression analyses (i.e., only the module was found to have a statistically significant impact on the critical incident counts).

4.3.1.3 Verbal frustration counts. Verbal expressions of frustration were identified from think aloud data upon review of OpenVULab screencasts. The total number of utterances of verbal frustration by each student for each module was recorded, and Figure 9 presents a histogram of the verbal frustration count data.

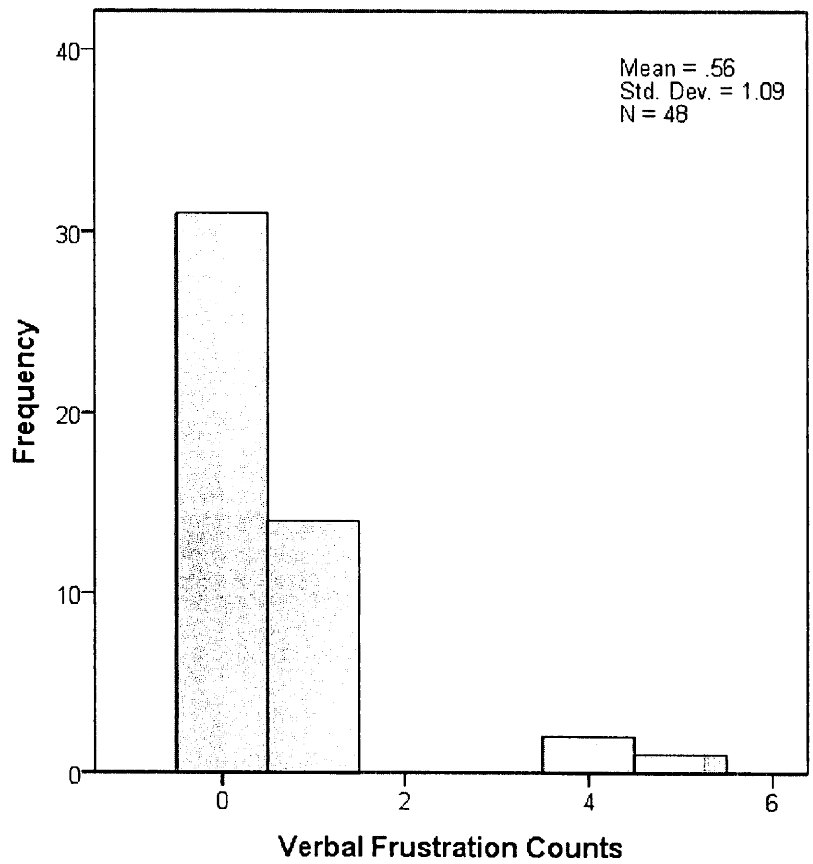


Figure 9. Histogram illustrating verbal frustration counts recorded during module completion

Figure 9 illustrates that there were a large number of zero counts for verbal frustration (i.e., many students did not verbally express frustration), and that the most common positive count value was one (i.e.., if students did verbally express frustration, they were most likely to do so only once during completion of a module). An exception to these trends was noted for three students – Julie, Janna, and Allison – who expressed verbal frustration four or five times during completion of the Scholarly Resources module, which they all completed in unmoderated sessions. Review of exit interview transcripts with these participants revealed that interviews conducted with Janna and Allison were lengthier than most as both students responded at length to several of the questions. Review of field notes made immediately after sessions with these students revealed that Allison participated in the study in the evening immediately following a class and that she appeared to be fatigued. It is possible that these characteristics of Janna and Allison (they were talkative students) and Allison's possible fatigue may have contributed to the higher than average verbal frustration counts. Review of these additional sources of data from Julie's session did not lend additional insight into why she may have verbally expressed frustration more than most of the participants.

It is possible that once a certain threshold of frustration was reached, these students were more prone to verbally expressing frustration many times. The mean number of critical incidents counted for individual students during completion of the Scholarly Resources module was 2.8 and the mode was 2; and counts of 5, 4, and 6 were recorded from review of Julie, Janna, and Allison's screencasts, respectively, which may support this notion. At the same time, the four other students whose critical incident counts for the Scholarly Resources module were above the mean (three of whom completed the Scholarly Resources module in a moderated session) expressed frustration either once or not at all, which may indicate that some students were simply less likely to verbalize frustration (particularly in a moderated session) and/or did not become as frustrated when encountering challenges.

The histogram of the frequency of verbal frustration counts presented in Figure 9 suggested that the data may be described as fitting a Poisson or negative binomial distribution (the mean count was found to be 0.54 and the variance was found to be 1.19), and regression analyses based on both of these distributions were conducted and the diagnostics were compared to determine which test may be most appropriate. In both analyses, the format of the testing session (moderated or unmoderated) and module (Scholarly or Credible Resources) were assigned as independent variables and verbal frustration counts as the dependent variable. Based on a comparison of the CAIC values generated from the regression analysis, the negative binomial regression analysis may have been more appropriate (CAIC of 103.311 compared to 106.203 from the Poisson regression). Regardless of whether a negative binomial or Poisson regression analysis was conducted, the format of testing was found to have a significant impact on the counts of verbal frustration (Wald's chi-square = 4.403 and $p = 0.036$ from the negative binomial regression; and Wald's chi-square = 6.779 and $p = 0.009$ from the Poisson regression) with higher counts tabulated from the unmoderated sessions. However, an examination of the regression diagnostics revealed that the standardized residuals did not fit very well to a normal distribution when either analysis was conducted. This can be an indication that the data were not ideally suited to either test and that it may not be appropriate to place a lot of weighting on the results.

The data were also collapsed into binary categories (zero counts and greater than zero counts) to allow for logistic regression analysis with the idea that students may either be prone to verbalize (a little or a lot) or not at all. In contrast to the Poisson and negative binomial regression analyses, logistic regression analysis suggested that only the module may have a borderline significant impact on the verbal frustration counts, with more counts arising from the Scholarly Resources module (Wald's chi-square = 3.267; p = 0.071) rather than the format of the testing session. Here again, the regression diagnostics were concerning as the standardized residuals did not appear to exhibit normality.

Overall, it is not entirely clear from statistical analyses alone whether or not the testing format and/or module had a significant impact on the verbal frustration counts. However, when the results of these analyses are combined with other relevant data from the study, they do appear plausible. It would not be surprising to find that there were higher counts of verbal frustration noted when students completed the Scholarly Resources module compared to the Credible Resources module. This is because students reported that this module was more challenging (refer to section *4.2.2.2 Post-session questionnaires*) and there were more critical incidents recorded when students completed this module (refer to section *4.3.1.2 Critical incident counts*). Moreover, data from exit interviews also supports the idea that the format of testing session could affect the amount of verbal frustration data. When students compared their comfort level with the think aloud protocol during moderated vs. unmoderated sessions, many students indicated that they felt more comfortable thinking aloud (possibly including verbally expressing frustration) when working alone in the testing room (refer to section *4.3.2.2 Exit interview data related to moderated vs. unmoderated sessions*).

4.3.2 Student perceptions and preferences of moderated vs. unmoderated sessions. Data from post-test questionnaires and exit interviews helped to shed more light than the quantitative analyses on how the different formats of testing sessions (moderated vs. unmoderated) may have impacted the students' perceptions of accessibility of the modules as well as their thoughts on participating in the study.

4.3.2.1 Post-test questionnaire data from moderated vs. unmoderated sessions. Tables 12 and 13 report the post-test questionnaire results for the Scholarly and Credible Resources modules, respectively, broken down according to the format of the testing session (i.e., moderated vs. unmoderated). Mann-Whitney U comparison testing was performed on these data to determine whether the format of the testing session affected the students' perceptions of the difficulty of the module and/or ease or comfort level with participating in the session.

Self-reported ease and comfort level with completion of the Scholarly Resources module was not significantly affected by the format of the testing session ($U = 50.5$ and $p = 0.219$ for overall ease of completing the module; $U = 50$ and $p = 0.219$ for ease of participating in the session; and $U = 71$ and $p = 0.977$ for comfort level participating in the session). Ten of the post-test questions pertained to ease of specific aspects of module completion. Of these ten questions, results of the Mann-Whitney U testing were interesting for the questions "How easy was it to understand the video included within this module?" and "How easy was it to work with websites outside of Moodle while completing this module?" with $U = 33$ and $p = 0.024$; and $U = 37$ and $p = 0.045$, respectively, with greater ease reported following participation in a moderated session. For a visual depiction of how the responses to these questions differed across the two testing formats, refer also to Figure 10, which presents these data in the form of clustered bar charts.

Table 12.

Participant Feedback from Questionnaires Administered After Completion of the Scholarly Resources Module under Different Testing Conditions

Question	Mean		Median		Mode		Range	
	U	M	U	M	U	M	U	M
Overall, how easy was it to work within Moodle while completing this module?	3.5	3.9	4	4	4	4	3	2
How easy was it to read the instructions included within this module?	3.4	4.0	3.5	4	3[a]	4	4	2
How easy was it to understand the instructions included within this module?	3.3	3.8	3	4	3	4	3	3
How easy was it to access the PowerPoint presentation within this module?	4.3	4.6	4	5	4	5	3	2
How easy was it to understand the PowerPoint presentation within this module?	3.9	4.0	4	4	4	3[a]	3	2
How easy was it to access the video included within this module?	4.2	4.5	4	5	4	5	3	2

How easy was it to understand the video included within this module?	3.3	4.3	4	4.5	4[a]	5	3	3
How easy was the question included at the end of this module?	3.5	4.1	4	4	4	4	4	2
How easy was it to post your answer to the question on the discussion forum in Moodle?	4.0	4.5	4	5	4	5	2	2
How easy was it to work with websites outside of Moodle while completing this module?	2.9	3.9	3	4	3	5	3	3
Overall, how easy was it to complete this module?	3.3	3.8	3	4	3	4	3	3
Overall, how easy was it for you to participate in this testing session?	3.7	4.2	3.5	4	3[a]	4	3	2
Overall, how would you rate your comfort level with participating in this session?	3.8	3.8	4	4	3	3[a]	3	2

Note. Survey questions were Likert-scaled, with 1 = very difficult and 5 = very easy for questions beginning with "How easy was it…" and 1 = very low and 5 = very high for the question beginning with "Overall, how would you rate your comfort level…" M = moderated testing session; U = unmoderated testing session.

[a]Multiple modes exist. The smallest value is shown.

Table 13.

Participant Feedback from Questionnaires Administered After Completion of the Credible Resources Module under Different Testing Conditions

Question	Mean		Median		Mode		Range	
	U	M	U	M	U	M	U	M
Overall, how easy was it to work within Moodle while completing this module?	4.5	3.8	4.5	4	4[a]	4	1	3
How easy was it to read the instructions included within this module?	4.3	3.4	4	3	4	3	2	4
How easy was it to understand the instructions included within this module?	4.3	3.9	4	4	4	4	2	3
How easy was it to access the PowerPoint presentation within this module?	4.6	4.3	5	4.5	5	5	1	2
How easy was it to understand the PowerPoint presentation within this module?	4.2	3.9	4	4	4	4	2	4
How easy was it to access the video included within this module?	4.5	4.3	4.5	4.5	4[a]	5	1	3

How easy was it to understand the video included within this module?	4.6	3.8	5	4	5	4	1	3
How easy was the question included at the end of this module?	4.5	3.9	4.5	4	4	4	1	2
How easy was it to post your answer to the question on the discussion forum in Moodle?	4.6	4.2	5	4.5	5	5	1	3
How easy was it to work with websites outside of Moodle while completing this module?	4.3	4.1	4	4	4	4	2	3
Overall, how easy was it to complete this module?	4.5	4.1	5	4	5	5	2	3
Overall, how easy was it for you to participate in this testing session?	4.6	3.7	5	4	5	3	2	4
Overall, how would you rate your comfort level with participating in this session?	4.3	3.4	4	3.5	4	4	2	2

Note. Survey questions were Likert-scaled, with 1 = very difficult and 5 = very easy for questions beginning with "How easy was it…" and 1 = very low and 5 = very high for the question beginning with "Overall, how would you rate your comfort level…" M = moderated testing session; U = unmoderated testing session.

[a]Multiple modes exist. The smallest value is shown.

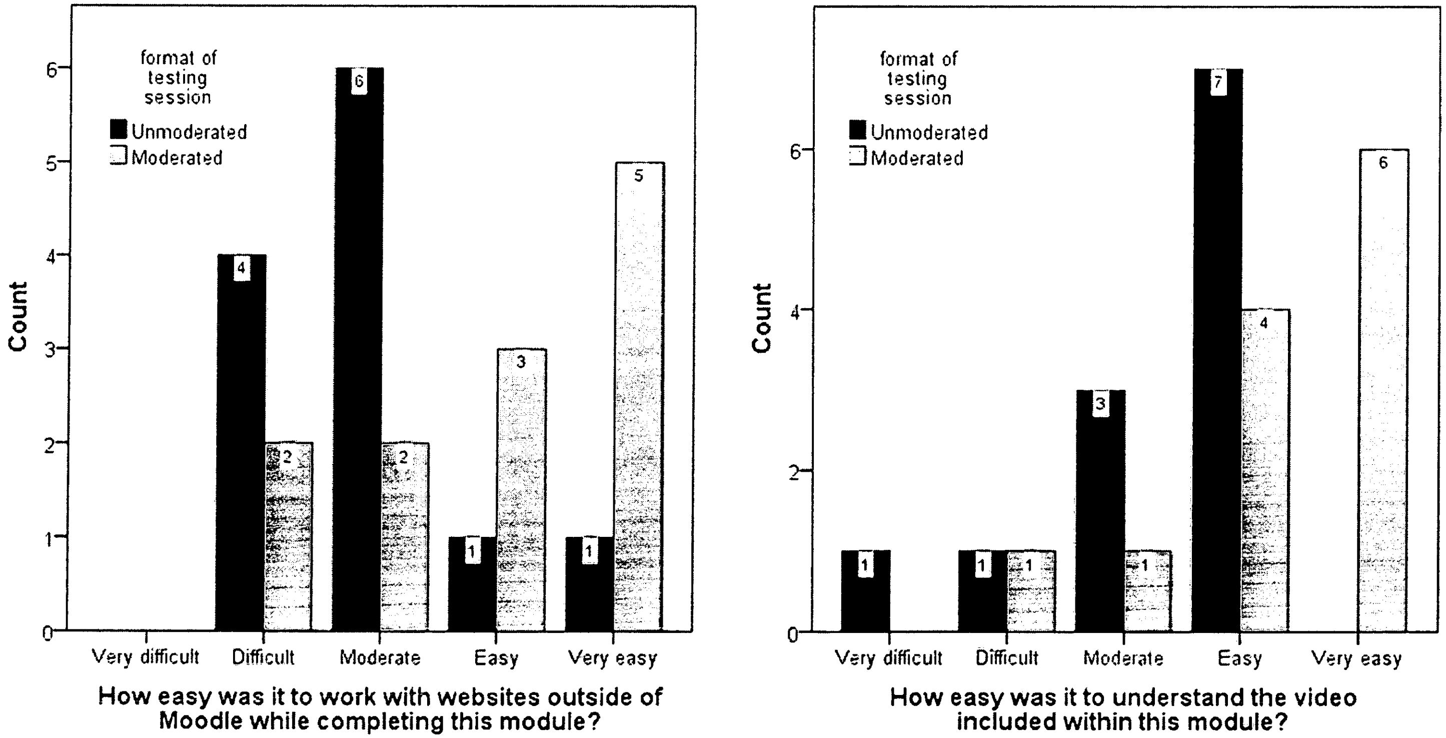
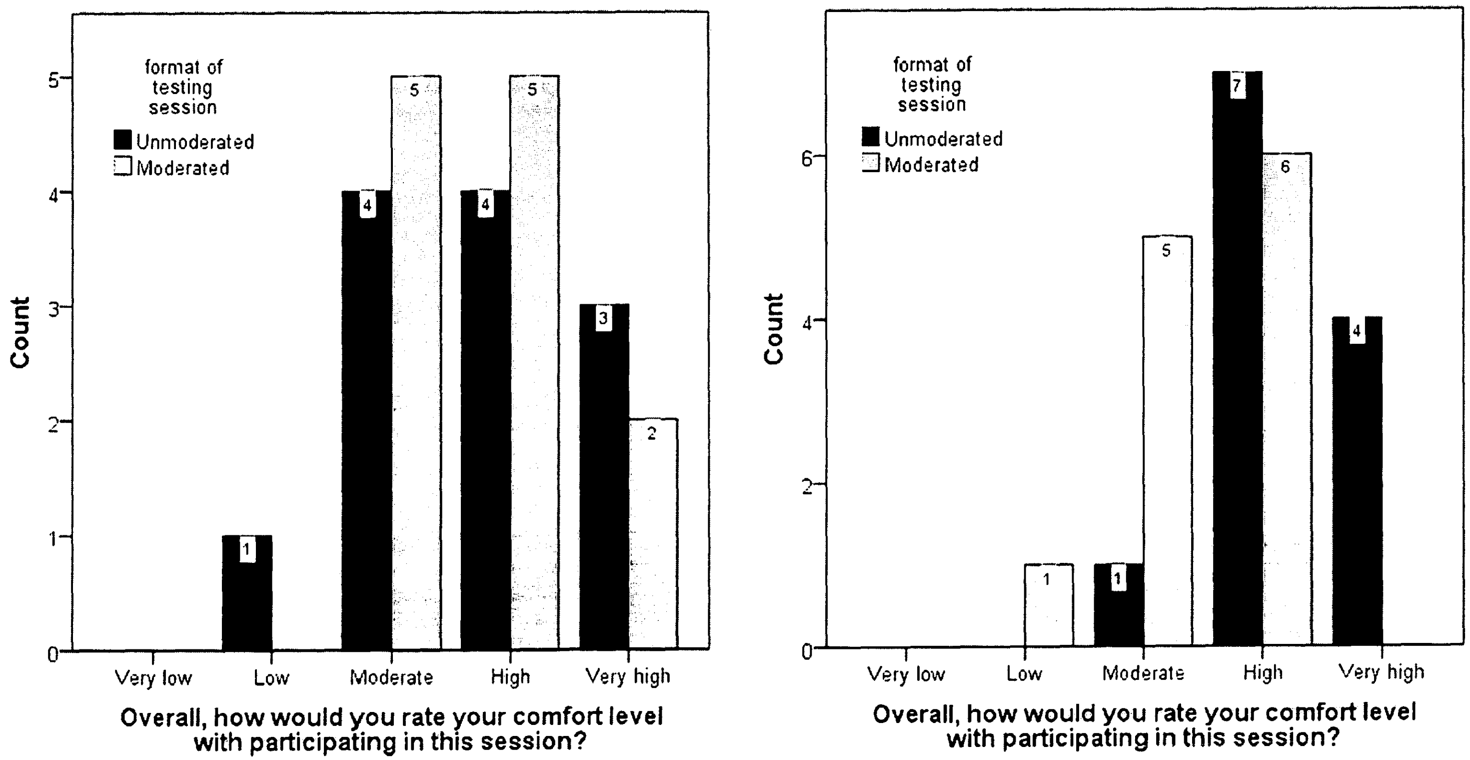


Figure 10. Frequency counts of student perceptions of the relative ease of completing aspects of the Scholarly Resources module in moderated compared to unmoderated testing sessions.

When the α is adjusted from 0.05 to 0.005 by the Bonferroni method (Bland, 2000; Elston & Johnson, 2008) to account for 10 individual comparisons, the differences in responses to these questions following moderated and unmoderated testing sessions are not deemed significantly different. It is interesting, however, that the YouTube video and websites external to Moodle were highlighted by this analysis, as these were aspects of the Scholarly Resources module that students were observed to encounter the most difficulty with and were highlighted as most challenging aspects by the students in exit interviews (refer to Table 10). It is possible that students rated these aspects of the module as being easier during the moderated session due to the presence of the researcher and the possibility of asking for help, though students were instructed to ask for help only if they encountered a major technical difficulty such as loss of Internet connection or computer malfunction. In instances where students did turn to ask for help, they were instructed to do what they would normally do if they had encountered this problem while working independently. Indeed, when asked about his experience in completing the Scholarly Resources module during a moderated session, David commented "I guess that I probably shouldn't have been asking you stuff. But when I was asking you stuff, you said 'just do whatever you'd regularly do normally' and I was like, oh, OK." When asked if that reassurance was helpful, David replied "Yeah, because I was like, what if this thing doesn't work? I was like, oh, what would I do at home. I would just try something else." Leanne also indicated that she felt reassured after receiving clarification that she could do what she would normally do during the testing session.

Several post-test questions from completion of the Credible Resources module were highlighted as interesting from Mann-Whitney U testing. Both ease of participating and comfort level participating were reported as higher following completion of this module in an unmoderated session ($U = 36$ and $p = 0.039$ for ease participating in the session; $U = 29.5$ and $p = 0.012$ for comfort level participating in the session). Similarly, students reported increased ease with working within Moodle, reading instructions, understanding the video, and answering the question within this module when they worked on the module during unmoderated sessions (however, none of these results are deemed significant when the α is adjusted from 0.05 to 0.005 by the Bonferroni method to account for the 10 individual comparisons). It is possible that since students found this module to be relatively easy overall, they did not take comfort in having a researcher present in the room with them and preferred to work independently. In other words, preference and increased comfort level with working independently may have exceeded the perceived benefit of having direct access to the researcher when this module was completed. For a visual depiction of this trend, refer to Figure 11. The following section, *4.3.2.2 Exit interview data related to moderated vs. unmoderated sessions*, provides an overview of students' thoughts on their preferences for working independently vs. with a researcher present, and provides additional supporting evidence for the ideas put forth here.



a) Scholarly Resources module

b) Credible Resources module

Figure 11. Frequency counts of student perceptions of the relative comfort in participating in accessibility testing sessions in moderated compared to unmoderated formats.

4.3.2.2 Exit interview data related to moderated vs. unmoderated sessions. When asked during exit interviews, the majority of the participants in this study ($N = 15$; 62.5%) indicated that they preferred to participate in the unmoderated portion of the testing session. Of the remaining students, six (25%) indicated that they had no preference for moderated vs. unmoderated sessions, while three (12.5%) students – Allan, Donna, and Susan (all three of whom identified as persons with learning disabilities, with Susan also disclosing a mental health disability) – indicated that if they were to participate again they would prefer to participate in a moderated session. For Allan and Donna, the moderated testing format was preferred because they had access to help from the researcher if needed. For example, even though Allan felt that he was able to complete both modules without a lot of difficulty, when asked about what he liked best about the moderated session, he responded:

> I guess just knowing that you were in the room so that if there were any problems then I know that you would be in here and you could kind of guide me through any problems. But, um, you know luckily there were no problems.

Indeed, many students (including those who preferred the moderated session) did concede that a drawback of participating in the unmoderated session was the lack of immediate access to help should the need arise.

Susan, on the other hand, indicated that she felt a lot of uncertainty about what to expect during the testing sessions and that was the primary reason that she felt more comfortable in the presence of a researcher. She went on to say that:

> Umm, if it [a remote testing session] was from home I wouldn't have access to anyone. I wouldn't know, I wouldn't be confident or I wouldn't know if I'm doing it right. If I'm missing something. I'd be worried if I'm missing something, if I skipped something really important I wouldn't know about it and stuff.

She also mentioned that she may not mind participating again in an unmoderated session provided that there was some sort of orientation prior to that where she could ask questions about what to expect.

In addition to valuing access to help during testing sessions, students' comments about the format of the testing sessions also touched on the themes of engagement and comfort level, which are discussed in more detail below.

4.3.2.2.1 Engagement. For some students, including Mandy, Eva, and Karen, it was easier to remain engaged with the modules when working in the testing room alone (during unmoderated testing) because the presence of a researcher in the room could be distracting. According to Mandy, "Umm, the best [part about the unmoderated session] is that I got to just, you know, focus on the computer. My attention was all on that." Similarly, Karen stated that:

> I think, for me – because I personally prefer to have a quiet setting, even if I'm writing my tests or even if I have someone around – I just like the fact that I know I'm by myself and I don't have to look and get distracted. So there is a lot of distraction knowing someone is there.

Eva also appreciated that she could simply focus on the computer when she was working alone.

Ollie, on the other hand, when asked what she liked the best about participating in the moderated testing session said "I was more engaged. Like, I was definitely more engaged." While Jeremy did not use the term engagement specifically in discussing a drawback of unmoderated testing, he did imply that he could be more distracted when participating in an unmoderated session in that he may be more likely to procrastinate.

Several students indicated that working alone in an unmoderated session better suited their preferred working or learning style compared to the moderated format. Aliya, Mandy, and Eva all mentioned that they like to work alone better. For Aliya, listening to music with earbud headphones while she works helps her to concentrate. During the moderated testing session that she participated in, she removed one earbud so that she would be able to interact with the researcher if necessary and she found that it was more difficult for her to concentrate during that session.

4.3.2.2.2 Comfort level. Several students reported feeling self-conscious during the moderated session and that they were more apt to act "naturally" during the unmoderated session. Leanne summed up her thoughts on this by saying:

> Well I always get a little anxious if someone is watching me in terms of like, typing or like navigating. I just want to make sure that I do a good job and it's different like to talk out loud.

Connie reported feeling a lot of anxiety during the moderated session. She remarked:

> Oh, I was just nervous, like nerve-racking because I was knowing that you were right behind me. And I was like, oh, how am I supposed to do this? Is this going to be a trick? Is there going to be some sort of like hidden problem?

Jessica also felt more comfortable to encounter problems and make mistakes while working alone as she said "if I made some mistake or I turned back on my question, you would only see it in the [screencast] video. (laughs) You wouldn't see it at the moment that I do a mistake."

The use of the video camera to record the student working at the computer during moderated sessions was also frequently discussed by students. Allison mentioned that she was smiling at times because she felt self-conscious about being videotaped and indeed it was observed that she smiled throughout most of the moderated session, making it difficult to glean additional insight from the video camera recording about her frustration level. Similarly, Jessica was also observed to smile each time she spoke aloud during the moderated session. For Karen, the presence of the video camera in the testing room was a distraction for her even during the unmoderated portion of her session when it was not turned on. She felt that the presence of the video camera in the room was the least positive aspect of participating in the unmoderated session as "I kept looking at it. Even though it wasn't on. It was just there. It was the most omnipresent part." Brad declined to be videotaped at all, as he expressed concern about the creation of a digital video file even after being reassured that the video would not be shared with others and would be deleted at the conclusion of the study. Overall, there were a range of responses to the presence of the video camera as for some students, the camera made them feel self-conscious throughout the entire moderated session, while other students reported feeling self-conscious at first but soon forgot about the video camera as they became more engaged with the module.

When reflecting on their experience during the unmoderated session, many students described feeling that they acted more naturally when working alone. According to Morgan, "I guess I could really pretend I was, you know, on my own. At home or something. Umm, yeah, it's just a calm, quiet environment." Several students commented that they felt more free to explore on the computer when working alone. According to Shelley:

> I felt that, like, I can really, I don't know, like play around with the buttons more. Um, just because I'm not being watched? At the same time I was comfortable before because it is a research study, but I think that when someone is not there you kind of let down a guard so you can kind of do what you would at home more.

Donna indicated that she would have felt more at ease to express frustration via her body language and think aloud verbalizations if she had been at home. She stated:

> There were parts where I caught myself wanting to sit back and at home I would go, Oh! And I would have been very dramatic about the whole event.... there were a few times where, um, especially in the first one, I was really starting to get frustrated because I had no clue. So at home I would've probably sat back and said, this is ridiculous! (laughing)

Interestingly, Donna was found to encounter the highest number of critical incidents compared to all other students, yet did not verbally express her frustration even once.

Donna was not the only student who discussed the think aloud protocol when reflecting back on their experience participating in the study. Morgan said "Well I felt crazy (laughs) talking to myself. I don't think anyone talks out loud that much."

Similarly, Janna commented:

> I wasn't trying to, like I didn't feel like I was trying to suddenly impress you when you were in the room. I didn't feel like totally, uh, but it felt easier to – when I was by myself – to keep on talking to myself. I didn't feel like I was a crazy person.

Julie mentioned that she felt that her think aloud was a little bit more "forced" during the moderated session compared to the unmoderated session. However, it was harder for her to remember to vocalize her thoughts when she was working alone as she felt more aware of the expectation to think aloud during the moderated session. For Ollie, it was difficult to decide which format of testing session she preferred – moderated or unmoderated – because while she felt more engaged participating in the moderated session, she felt more self-conscious thinking out loud when she was not alone.

4.3.3 Additional data from moderated sessions. The video camera recordings were created only during the moderated sessions and, if the unmoderated sessions had taken place in a setting where the student and researcher did not interact in real time (e.g., as may take place in remote accessibility testing), exit interviews may not have taken place. During the reporting of results of Research Objective 1 (refer to section *4.2 Research Objective 1*) where objective accessibility testing was compared to subjective testing, the added value of the video recordings and exit interviews were discussed in sections *4.2.2.1 OpenVULab screencasts and video camera recordings* and section *4.2.2.3 Exit interviews*. Overall, the video camera recordings were not found to provide very much useful information that was not already evident from other sources of data (e.g., OpenVULab screencasts and think aloud verbalizations) and many students also described feeling self-conscious in the presence of the video camera, which may have further reduced the usefulness of the recordings (e.g., when students smiled frequently as a result). Exit interviews, however, yielded much valuable information about the most and least challenging aspects of each module, which served to both reinforce findings regarding accessibility problems from other sources of data, and provide new insight about factors that positively influenced accessibility from the students' perspective.

4.3.4 Summary of findings from Research Objective 2. Statistical analyses comparing the data from moderated and unmoderated accessibility testing sessions suggested that the format of the testing session did not significantly impact the efficiency of module completion or the critical incident counts (though there were more critical incidents observed as students completed the Scholarly Resources module). It is unclear whether verbal frustration counts were affected by the format of testing as the results of regression analysis point to a possibility of significantly increased counts of verbal frustration from the unmoderated sessions, though the regression diagnostics are troubling as examination of the standardized residuals suggest that the model may not be specified correctly and the results therefore may not be meaningful. Similarly, results of logistic regression analysis are suggestive of increased counts of verbal frustration during completion of the Scholarly Resources module but it is unclear whether the results of the regression analysis are meaningful. However, if combined with exit interview comments related to the different testing formats and the relative difficulty of the Scholarly

Resources module, it is not unreasonable to expect that students may have felt more comfortable verbally expressing frustration during the unmoderated sessions as many student felt less self-conscious while working alone, and that there may have been more instances of verbal frustration expressed during completion of the more challenging module.

Comparison of responses from post-test questionnaires following completion of each module under the different testing conditions suggested that students may have found the challenging components of the Scholarly Resources module to be easier to handle during the moderated sessions where they had access to the researcher. Though help was not provided, some students expressed that they took comfort in knowing that they could ask for help if needed and that they were reassured when told to proceed to work through problems as they normally would if they were working alone. In contrast, students reported that the overall ease of Credible Resources module completion, as well as specific aspects of the module and their comfort level with participating in the session, was greater during the unmoderated sessions. This may be a reflection of the fact that most students reported preferring the unmoderated format of testing where they felt less self-conscious and more free to act naturally as they worked on the computer. This preference for unmoderated testing may have over-ridden value associated with access to help when completing the less challenging Credible Resources module.

Engagement and comfort level were frequently discussed by students as they compared their perceptions of participating in moderated vs. unmoderated testing sessions. While not all students were in agreement about which setting led to the highest level of engagement and comfort level, the general consensus was that most students preferred to work alone and felt more comfortable interacting with the computer and thinking aloud when they were alone. Moreover, many students felt particularly uncomfortable with the presence of the video camera during the moderated sessions. Given that the video camera recordings provided little useful information, it is possible that the discomfort that the camera caused the students and the potential impact of the camera on students' behavior may have outweighed benefits of using it.

CHAPTER 5: DISCUSSION AND CONCLUSIONS

5.1 Objective vs. Subjective E-Learning Accessibility Testing

The extent to which objective measures of the accessibility of e-learning technologies (using automated tools) were able to predict the subjective accessibility experience of students as they completed modules of a mock online course was determined. Comparison of the results obtained from objective e-learning accessibility testing and subjective e-learning accessibility testing revealed that both methods of accessibility evaluation detect barriers to accessibility. However, the two methods of accessibility evaluation provided different information about the accessibility of the same e-learning infrastructure. Moreover, the subjective experiences of individual students in this study were different even though they interacted with the same e-learning technologies. As such, these findings indicate that the degree to which e-learning infrastructure will be accessible to students, and the nature of accessibility barriers encountered, will vary within a student population.

These findings highlight the importance of taking a multi-faceted approach to e-learning accessibility evaluation, by employing a variety of methods of accessibility evaluation in order to fully explore how to maximize accessibility for all students. Relatively few accessibility problems were predicted or observed for the e-learning infrastructure included in this study. However, the results of this study have demonstrated that there would be a benefit to many students (including students with and without disabilities) to consider how to enhance the accessibility of even e-learning technologies that could be rated as highly accessible overall.

5.1.1 Strengths and shortcomings of objective and subjective accessibility testing. When used to evaluate e-learning infrastructure included in this study, automated tools were effective in identifying several potential accessibility problems that may make it difficult for students who use assistive technologies to access, interact with, and/or understand all content therein. In addition, correction of potential problems such as increasing text-to-background color contrast and removing duplicate PowerPoint slide titles may increase accessibility of some of the infrastructure for students with print disabilities as well as students without disabilities. While these identified potential problems were not observed to negatively affect the participants in this study, correcting these problems could be beneficial to other students and would help towards achieving conformance with the WCAG 2.0.

Vigo and Brajnik (2011) have described several advantages of automated accessibility evaluation, including the ease of generating a list of potential problems using automated tools, and the affordability of using them (e.g., vs. recruiting and compensating users or expert evaluators for their time). These advantages of objective accessibility testing as compared to subjective accessibility testing were noted in this study. However, several additional disadvantages were also evident. For example, one drawback to using automated tools in this study was the challenge in interpreting the output. In other words, the accessibility (as it relates to understandability) of the results of objective testing may be rated low by researchers or practitioners who are not experts in interpretation of the results. Expertise in computer programming, for example, would help to increase the ease of understanding the results of objective testing. However, as automated testing tools are only as effective as the pre-set criteria (e.g., web accessibility guidelines) that they are designed to evaluate against (Vigo & Brajnik, 2011), even if the output of the tools is understood by an expert who is using them, limitations of the underlying criteria may reduce their effectiveness. Recall from section *2.5.2.1 Limitations of the WCAG 2.0* that concern has been raised about relying heavily on conformance testing to assess accessibility given that this may not effectively pinpoint some types of accessibility or usability problems (Brajnik, et al., 2012; Kapsi, et al., 2009b; Kelly, et al., 2005; Ribera, et al., 2009). Indeed, the automated tools used in this study were not effective in predicting the accessibility problems encountered by the study participants.

In their evaluation of several automated accessibility evaluation tools, Vigo and Brajnik (2011) note that the validity of a given tool (i.e., the ability of the tool to truly evaluate the accessibility of a website) can be distinguished as validity with respect to conformance (i.e., how well the tool evaluates a website against Web accessibility guidelines) and validity with respect to accessibility in use (i.e., how well the tool predicts the subjective accessibility experience of users). A drawback to validity of automated tools as related to validity with respect to accessibility in use is the high proportion of false positives and false negatives that they report (Brajnik, 2004; Vigo & Brajnik, 2011). While false positives were not examined for in this study, the inability of the tools used to predict the accessibility problems that the study participants encountered supports previous research which has found that automated tools are prone to high proportions of false negatives (i.e., failing to report true accessibility problems). According to Brajnik

(2004), the only way to deal with false negatives is to employ user-centered (subjective) testing.

Subjective testing with student participants shed light on the accessibility of the mock online course by allowing for the preparation of a list of observed accessibility problems, and providing insight into the students' perceptions of accessibility of each module and the e-learning infrastructure therein. In this manner, not only were actual accessibility problems identified, the impact of the problems was also revealed. This is in contrast to output from automated tools, which does not necessarily indicate the severity of the predicted problems. Particularly troublesome were problems related to accessing an understanding of module content. According to Bohman and Anderson (2005), the design of computer algorithms to support automated testing of understandability (e.g., evaluating for "clear and simple" text) is difficult if not theoretically impossible. In this study, efforts were made to include simple and understandable text within the e-learning modules, though it was not until students were observed engaging with the Scholarly Resources module that the confusing nature of some of the module instructions was revealed.

Another key way in which the data from objective and subjective testing differed in this study was that the subjective testing revealed not only disabling attributes of the e-learning infrastructure, but enabling attributes as well. As Vigo and Brajnik have noted, "accessibility, like usability, is much easier to be noticed when it is missing" (2011, p. 151). By describing enabling attributes of the modules during exit interviews, students in this study highlighted the enabling effects of attributes such as familiarity, simplicity, enhanced learner control, enhanced tolerance for learner differences, and social presence.

Not only did this offer insight into what students found favorable about the modules, these data also provided insight into how to effectively improve aspects of the modules that students found less favorable. Moreover, these same data may be applied to proactive efforts to design accessibility into e-learning infrastructure in other contexts, and may therefore transcend a particular accessibility evaluation study.

Finally, a third difference in objective vs. subjective accessibility evaluation was that objective evaluation examines a particular e-learning technology in isolation – rather than considering the context of use. Consider the following example to illustrate this point. The confusing instructions within the Scholarly Resources module read as follows:

1. View a Virginia Commonwealth University (VCU) library video demonstrating how to use the online catalog.
2. Use the VCU library catalogue to access the journal titled Academic Leadership.
3. Skim the abstract of an article in the current issue of the journal for interest.

There were not any predicted problems with this text highlighted by automated tools used in the study to evaluate the accessibility of the webpage that the text was included on (and, as an aside, the use of the automated tool http://www.read-able.com/check.php to evaluate the readability of this excerpt of text reports that it should be "easily understood by 14 to 15 year olds"). However, this represented a stumbling point for many of the student participants. The problem with this text was that, rather than completing each of the three steps in turn, some students attempted to merge steps two and three to look for an abstract from the journal directly from the library catalogue. The intended process was for the students to use the library catalogue to link to the journal website and, once on the journal website, to read an abstract from a recent article. The context in which the students read and followed the instructions caused confusion, as the focus on use of the library catalogue prompted students to expect that they should use the library catalogue to complete both steps two and three. The importance of considering the learning context when assessing e-learning accessibility is discussed further in the following section.

5.1.2 Alternative approaches to conceptualizing accessibility evaluation. To evaluate e-learning accessibility, Kelly, Phipps, and Swift (2004) advocate looking beyond guideline conformance to also take into account issues related to pedagogy, the learning context, and learning styles. In describing a holistic approach to web accessibility (Kelly, et al., 2004), they have stated that "e-learning is a process, not an event" (p.8) and "accessibility is primarily about people and not about technologies" (p. 11). Similarly, Sloan et al. (2006) have suggested that "contextual web accessibility" includes consideration of user characteristics, desired functionality of a website (i.e., activities that the website should support), technological requirements (e.g., software or plug-ins needed to use the site), and performance requirements (e.g., comprehension outcomes). Broadening the perspective on e-learning accessibility and accessibility evaluation to consider these contextual factors may therefore help to anticipate and uncover potential accessibility problems (or solutions) that actual users of e-learning infrastructure may encounter. To this end, Cooper, Sloan, Kelly, and Lewthwaite (2012) advocate moving from evaluating accessibility of individual learning resources towards evaluating the accessibility of learning outcomes.

When the notion of accessibility of learning outcomes (vs. learning resources) is considered in relation to the results of this study, the focus of accessibility evaluation would shift from evaluating individual web pages and content therein to considering whether students would be provided with sufficient means to meet the learning outcomes stated at the beginning of each module. This more holistic approach to accessibility evaluation will more effectively take into account learner diversity, as there will be multiple means available for students to meet the same learning outcomes. So long as each learner is able to identify one or more approaches that can be taken to meet the learning outcomes that is accessible for that individual, the e-learning environment would be accessible for that population of learners. While conformance of individual components of e-learning infrastructure to web accessibility guidelines such as the WCAG 2.0 may be a demonstrable way to meet accessibility legislation such as the AODA (Sloan & Kelly, 2008), the results of this study point to the value of considering learner diversity and flexibility of the learning environment as much as possible to truly increase the subjective accessibility of e-learning. This notion of separating learning methods from learning outcomes and thus accounting for learner diversity is aligned with the Universal Design for Learning framework, which is discussed next.

5.1.2.1 Universal Design for Learning and accessibility. The Universal Design for Learning (UDL) framework as described by the US non-profit organization *Center for Applied Special Technology* (CAST; Rose & Meyer, 2002) was inspired by the Universal Design (UD) movement in architecture. The underlying premise of UD is to proactively design physical spaces so that they may be used by a wide variety of patrons, including those with and without disabilities (Center for Universal Design, 1997; Wilkoff & Abed, 1994). CAST's iteration of UDL draws both on the underlying premise of UD and the neuroscience of learning, with the overarching idea that the curriculum (rather than the students) should shoulder the burden of flexibility and adaptability (Rose & Meyer, 2002). The UDL framework is based on three overarching principles, namely to provide multiple means of representation, expression, and motivation. Associated with these three overarching principles are specific guidelines intended to serve as a guiding framework for instruction (CAST, 2011). This notion of the importance of a flexible learning environment towards meeting the needs of diverse learner populations is aligned with the way in which accessibility has been conceptualized in this study (refer to section 3.1 *Defining Digital Disability and E-Learning Accessibility*) and is supported by the findings of this study.

The UDL principle of providing multiple means of representation is intended to reflect learner differences across recognition learning networks (Rose & Meyer, 2002). Recognition networks are involved in the receipt and analysis of information – described by Rose and Meyer (2002) as the "what" of learning. In this study, there were several instances in which students indicated that they preferred and/or found it easier to learn from one form of representation of content over another. For example, while many students liked the simple single-page structure of the module homepages, Penny found that it was difficult and overwhelming to be presented with the entire module at once. To cite another example, Morgan finds that she learns well from text while David expressed concern that he may miss important information that is presented only in text form. The way in which material was represented affected the ability of the students to construct an understanding of the material, and it therefore stands to reason that e-learning environments that include multiple means of representation may be more widely accessible than those that do not. This assertion is true not only for students who identify as persons with disabilities (such as David, from these examples) but also for students who do not identify as being persons with disabilities (including Morgan and Penny, from these examples).

The UDL principle of providing multiple means of expression is intended to reflect learner differences across strategic learning networks (Rose & Meyer, 2002). Strategic learning networks are used when planning and executing actions – described by Rose and Meyer (2002) as the "how" part of learning. In this study, students were given quite prescriptive instructions within each module: The step-by-step module instructions asked students to take the same pathway towards learning and applying their learnings, and students were asked to describe this process in text form on a discussion forum. There were several problems that arose from this rigid module structure that were most pronounced during completion of the Scholarly Resources module. During completion of this module, students were asked to use one specific library catalogue to look for one specific resource, though the learning objectives for the module were more general (to locate a scholarly resource from an online library catalogue, and to access a current journal article). If the learning methods were separated from the learning outcomes for this module (as advocated by UDL), students could have been given the option of expressing their understanding of how to use a library catalogue of their choice and to locate any current journal article. It is possible that students may have chosen to work with e-learning infrastructure that they felt suited their needs best (e.g., which was familiar and/or simple, attributes that were revealed as themes positively impacting e-learning infrastructure accessibility). As a result, the accessibility of meeting the learning outcomes may have been increased. Similarly, students who feel more comfortable with or skilled at expressing themselves in non-written form may have appreciated the option of an alternate form of expressing their understanding such as by recording an audio message to complete the module. In this example of difficulty with the Scholarly Resources module, failure to successfully complete all aspects of the module due to difficulty with the prescribed learning methods was experienced by students with and without disabilities. As such, providing for multiple means of expression in e-learning environments could enhance accessibility for all students.

The UDL principle of providing multiple means of motivation is intended to reflect learner differences across affective learning networks (Rose & Meyer, 2002). Affective networks are used by learners when evaluating information and setting priorities – described by Rose and Meyer (2002) as the "why" of learning. In this study, engagement arose as a major theme that contributed positively to the ease and accessibility of e-learning, and difficulty remaining engaged was most commonly cited by participants as the most difficult part about completing the Credible Resources module. At the same time, study participants reported that they felt most engaged with different forms of e-learning infrastructure. For some students, the interactive nature of e-learning environments was itself highly motivating and engaging, while for others e-learning environments may provide more opportunity for distraction. These results strongly point to the need to work in multiple means of motivation into e-learning environments so that all students are maximally engaged and motivated to learn.

This discussion has shown that UDL (providing multiple means of representation, expression, and motivation) would be beneficial to all students in e-learning environments, regardless of whether or not they identify as persons with disabilities. While the presence or absence of self-reported disability did not appear to have a statistically significant impact on the data gleaned from the OpenVULab screencasts or in themes on e-learning accessibility identified from exit interview analysis, it is possible that there may have been significant differences noted in the experiences of the students with and without disabilities in a different context (e.g., with different modules) or if the students identified with different disabilities. However, the results of this study strongly highlight the importance of accessibility and the value of UDL for all students.

5.2 Moderated vs. Unmoderated E-Learning Accessibility Testing

Data obtained from moderated e-learning accessibility testing was compared to that obtained from unmoderated e-learning accessibility testing. In both testing methods, screen recording software recorded participants' on-screen interactions and verbalizations, though during moderated testing a researcher was present in the testing laboratory and a video camera was also used to record the student and workstation. Statistical analyses of the efficiency of module completion and critical incident count data indicated that the format of the testing session (moderated vs. unmoderated) did not significantly impact these data. The additional video camera data from the moderated sessions led to new insight in only a few instances where students were observed to be leaning forward to read small text within module e-learning infrastructure. At other times, however, the video camera data appeared to be unhelpful as some students were observed to be smiling during the session. It is less clear whether the verbal frustration counts were affected by the format of the testing session, though it is possible that students were more prone to verbally express frustration while completing a difficult module during an unmoderated session.

There were differences in student-reported data provided following completion of moderated and unmoderated sessions, as revealed by post-test questionnaire analyses. When completing a module that was deemed challenging by many students, specific problematic pieces of e-learning infrastructure were rated as easier to understand or work with when the module was completed in the presence of a researcher. However, when completing a module that was deemed less challenging overall, students instead rated several aspects of module completion as easier during the unmoderated session, where they also reported higher overall ease and comfort level with participation. Factors related to attributes of moderated and unmoderated testing sessions that may have influenced these data are discussed in the following section.

5.2.1 Factors that may influence student behavior in moderated vs. unmoderated conditions. The post-test questionnaire results suggest that students may take comfort in having access to help when they encounter very challenging technology, and that reassurance from a researcher may increase perceptions of accessibility of e-learning infrastructure. In addition, the possibility of researcher intervention can also be deemed positive from the researcher's perspective because the researcher has opportunity to exert more control over the testing session (e.g., re-directing students who deviate from the desired on-screen activities; reminding students to think aloud). However, when a researcher does intervene in an accessibility evaluation session, this may reduce the internal validity of the data, because a variable other than the independent variable (in this case, format of testing session) exerts influence on participant behavior.

This study also revealed that in instances where challenges encountered are not overly severe, students may find it easier and more comfortable to explore e-learning infrastructure when working alone. In this study, most students preferred participating in the unmoderated sessions because they felt more comfortable. In contrast, students reported feeling self-conscious in the presence of a researcher and video camera, and indicated that feelings of self-consciousness may have influenced their activities on the computer and/or their body language and verbalizations. Students reported feeling as though they could behave more "naturally" on the computer during the unmoderated session and that they could "let their guard down" and "click around more" when working alone. Most students who chose to reflect on the use of the think aloud protocol indicated that they felt more comfortable verbalizing their thoughts when they were alone in the testing classroom. Some students also speculated how their behavior may have differed if they had participated in the study from home. For example, one student commented that the frustration that she felt during part of the session would have been relayed in her body language and exclamations of frustration had she been working at home, though this was not evident in screencasts or the video camera recording from her participation in this study.

Taken together, these data suggest that low counts of verbal frustration identified from the screencasts in this study could reflect students' feelings of self-consciousness during the study, and that self-consciousness may have had a larger impact on these data during the moderated sessions. Students also reported feeling as though they behaved less naturally during the moderated sessions, which may also have affected the data. Overall, students may have felt somewhat more comfortable and natural during the unmoderated sessions, though still not as comfortable/natural as they may have felt if participating remotely from home. Indeed, this sentiment was expressed by students during exit interviews. This may explain in part why significant differences in data across the two formats of testing sessions were not observed for the efficiency and critical incident counts in this study: While students may have felt more comfortable during unmoderated sessions, both formats of sessions took place in an unnatural laboratory setting and thus it is possible that differences that may exist amongst data from moderated laboratory testing vs. unmoderated remote testing may be greater than what were observed in this study (which took place entirely in a laboratory setting). Further investigation could determine whether there is a gradient in terms of ecological validity of data when moving from a moderated session in a testing laboratory, to an unmoderated session in a testing laboratory, to an unmoderated testing session in the learners' natural working environment. The potential effects of researcher presence and study location and corresponding implications for e-learning accessibility evaluation are discussed further in the following section.

5.2.1.1 Naturalistic vs. laboratory studies. It has long been established that data may differ when collected in laboratory settings compared to "natural" field settings (Shaughnessy, Zechmeister, & Zechmeister, 2000; Wilson, 1977). For example, participants under observation in "artificial" laboratory settings may behave according to how they believe they are expected or allowed to behave in that environment (Wilson, 1977). This modification of behavior is what Bochner (1986) refers to as "contextual effects" (p. 167). The more naturally participants behave in a research study may be correlated with increased external validity because the findings of a study in which participants behave as they normally would (i.e., if they were not a part of the study) may be more generalizable to other settings (Shaughnessy, et al., 2000). The validity of e-learning accessibility testing data for infrastructure that students may be expected to engage with primarily outside of a formal classroom setting may therefore be enhanced if the study took place in the students' natural work environment (e.g., from home).

The obtrusiveness of observation is also believed to influence data. When participants are aware of being observed, they may either perform better (in a manner that they think will impress the researcher) or they may exhibit reduced performance due to anxiety (Williams, Klamen, & McGaghie, 2003). Even without purposeful intervention by the researcher, the potential for participant reactivity in response to researcher presence can therefore negatively impact internal validity (Campbell & Stanley, 1966). This "observer effect" has been reported in studies across various fields of inquiry, including, for example, education, psychology, criminology, and health sciences (for examples, refer to Guerico & Dixon, 2011; Masling & Stern, 1969; Spano, 2006; Yamamoto & Suzuki, 2012). In this study, students reported feeling as though they should stay on task as well as feelings of nervousness due to the presence of an observer, and therefore the obtrusiveness of observation may have influenced their behavior. Moreover, during unmoderated sessions (taking place in a laboratory setting as in this study, or even taking place remotely) there is the possibility that screencast creation alone could result in unwanted observer effects. For example, Tang and colleagues (2006) conducted a study in which study participants used screencast software to record collaborative computer usage. The participants worked with the software in the absence of a researcher and in their natural work environments (starting, pausing, and stopping the recording process themselves), though still felt that knowing that their on-screen interactions were being recorded influenced their behavior (Tang, et al., 2006). As such, while less obtrusive measures of e-learning accessibility evaluation may lead to more valid data, it is important to recognize that students may still not behave entirely naturally even when working from their natural working environment in the absence of a researcher.

5.3 Recommendations for E-Learning Accessibility Evaluation in Higher Education

The findings of this study have highlighted the value of conducting both objective and subjective e-learning accessibility testing in order to evaluate the accessibility of e-learning infrastructure used in higher education. Conformance testing of individual e-learning technologies against Web accessibility guidelines such as the WCAG 2.0 may assist institutions of higher education in technically meeting the requirements of non-discriminatory legislation, and, in doing so, will improve accessibility for some students in some contexts. However, this study has demonstrated that even WCAG 2.0 compliant e-learning technologies may not be accessible to all students in all contexts. To achieve e-learning accessibility for all individuals in a population of learners, a multi-faceted approach towards e-learning accessibility evaluation that includes a variety of stakeholders will be most effective.

To this end, the rainbow bridge metaphor for conceptualizing e-learning accessibility proposed by Seale (2006b) is very useful. She has proposed a modification of the traditional bridge metaphor for accessibility, in which a bridge connecting partial accessibility to optimal accessibility takes the form of a rainbow. As shown in Figure 12, the various stakeholders who may assist in building this bridge make up the different strands of the rainbow. According to Seale (2006b), these stakeholders may include students with disabilities, lecturers (faculty members), technologists (e.g., multimedia developers and IT services), support services (e.g., campus disability service providers), staff developers (who may provide professional development support to faculty), and senior managers (who may play a role in institution-wide policy development and implementation). The different colors of the strands of the rainbow are symbolic in several ways: Different colors represent diversity of viewpoints and different perspectives on accessibility. The coming together of the different strands to form the cohesive rainbow suggests that all stakeholders share responsibility for building the bridge and working towards optimal accessibility. In other words, the rainbow depicts both diversity and unity at the same time. The bridge itself may represent new practices or new knowledge that assists the institution in moving beyond partial accessibility towards optimal accessibility.

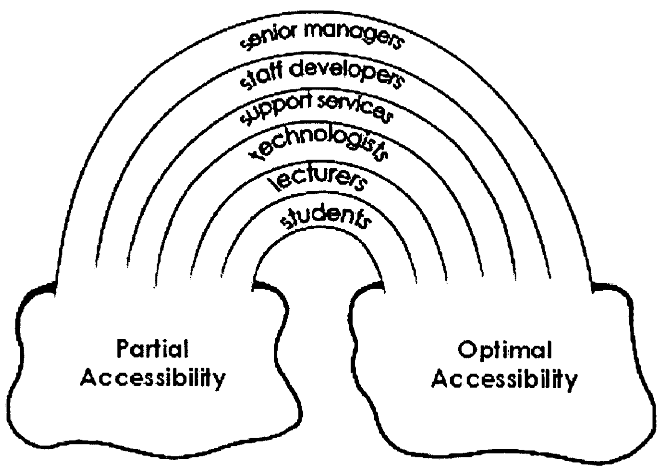


Figure 12. The rainbow bridge metaphor for conceptualizing accessibility. Image by J. K. Seale, retrieved from http://www.cjlt.ca/index.php/cjlt/article/view/56/53 Used under Creative Commons Attribution 3.0 License (http://creativecommons.org/licenses/by/3.0/).

When considered in concert with the findings of this study, the rainbow bridge metaphor can help towards articulating the necessary multi-faceted approach towards conducting e-learning accessibility evaluation. By way of illustration, consider the following hypothetical chain of events that could take place when a university is interested in procuring a new e-learning technology or evaluating technologies currently in use. Senior managers may implement policy to mandate that only e-learning technologies that have been evaluated for accessibility by objective measures (e.g., only those which are deemed WCAG 2.0 Level AA compliant) by the developers of the technology are considered for use by the university. On-campus learning technologists who are best qualified to interpret and respond to results of objective accessibility evaluation may also conduct objective accessibility evaluation of new and existing e-learning technologies on campus, consulting on-campus disability service providers as necessary to fully interpret the results and their implications. Learning technologists and disability services providers may also liaise with staff developers to relay important information regarding how to implement e-learning technologies in a maximally accessible fashion. This stage of the e-learning evaluation/implementation scheme may emphasize not only the accessibility of individual aspects of e-learning infrastructure, but the importance of an overarching holistic approach to accessibility which also considers the learning context and accessibility of learning outcomes. This insight can then be relayed to faculty with specific guidelines for how to use the e-learning technology in a manner that is most likely to be accessible for all individuals of diverse learner populations and aligned with the principles of UDL. Note that, while this illustration may suggest a linear unidirectional step-wise approach, there may be additional important interactions between various stakeholders that may take place and there is not necessarily one appropriate "direction" to the way in which interaction may take place across the different strands of the rainbow bridge.

In Seale's conceptualization of the rainbow bridge of accessibility, she includes students with disabilities as key stakeholders (Seale, 2006b). The findings of this study and the essence of the UDL framework emphasize that accessibility is important for all students. As such, all students (including those with and without disabilities) can offer important insight into the accessibility of e-learning infrastructure. Various stakeholders may collaborate to plan and conduct subjective accessibility evaluation of e-learning infrastructure. For example, faculty could partner with disability service providers, learning technologists, and students to design a testing scheme suitable for evaluating particular technologies. Either moderated or unmoderated subjective accessibility evaluation can be expected to provide valuable information about the true accessibility (accessibility in use) of the e-learning infrastructure. Factors such as the intended context of use of the e-learning infrastructure (e.g., fully online or face-to-face), the level of control over the accessibility evaluation session that the research team desires, and the needs and preferences of student participants may influence the choice of format of accessibility evaluation sessions. Ideally, results of e-learning accessibility evaluation would be disseminated to all key stakeholders of the institution (and beyond) so that the results may broadly inform the development of best practices regarding accessible implementation of e-learning infrastructure in higher education.

5.4 Future Work

Future investigations on methods of e-learning accessibility evaluation in higher education may build further on the research questions that have been posed here, and focus on new avenues of exploration that have emerged from this study.

In this study, students who identified as persons with learning disabilities (including one student who also disclosed mental health illness) and students who do not identify as persons with disabilities were included in subjective accessibility testing sessions. It was found that the experiences of these students with the e-learning infrastructure included in this study were similar regardless of whether or not they identify with disability in the traditional medical model sense. For example, similar amounts and types of critical incidents were encountered. Moreover, most and least challenging aspects of interacting with the e-learning infrastructure as well as general themes regarding factors that influence accessibility were also reported similarly by the two populations of students. It is not, however, clear from this study whether the same trends would hold true for different e-learning infrastructure (e.g., other forms of e-learning infrastructure, or other infrastructure that is deemed poorly accessible by objective evaluation methods) or if students who identify with different types of disabilities were included in such a study.

The literature does indicate that there is significant overlap between Web design principles developed to meet the needs of users with learning disabilities, physical disabilities, and visual impairment (Evett & Brown, 2005; Halbach, 2010; McCarthy & Swierenga, 2010). To further explore this in the context of e-learning accessibility evaluation, it would be instructive to compare results of subjective accessibility testing of the same e-learning infrastructure obtained from students who identify with different disabilities to determine to what extent their needs and preferences compare. This type of investigation may be helpful to ascertain which collection of student characteristics within an accessibility evaluation tester pool may be expected to yield the richest and most comprehensive set of data regarding accessibility of e-learning technologies. This information could aid in the design of subjective accessibility evaluation studies, by shedding more light on the optimal number and characteristics of testers that may be required to maximize external (population) validity and thus generalizability of results to the broader post-secondary population.

While it can be difficult to recruit student participants for research studies, let alone participants with specific characteristics (e.g., as related to ability/disability), another advantage to carrying out additional studies with a larger sample size is to increase the power of the statistical analyses. A larger sample size is generally associated with greater statistical power, which is a measure of the ability of statistical analysis to detect an effect of an independent variable (D'Agostino, Sullivan, & Beiser, 2006). In this context, for example, additional investigation with a larger number of student participants may be better able to determine whether verbal frustration counts are significantly affected by the format of a testing session and/or the e-learning infrastructure being tested. As such, greater sample size and greater statistical power may increase the internal validity of reported results such that more firm conclusions may be drawn from quantitative analyses.

Another area that warrants further investigation is examining different methods of unmoderated accessibility evaluation. There are several additional related research questions that could be asked, including: How do data from moderated (laboratory setting), unmoderated (laboratory setting), and remote (natural setting) e-learning accessibility evaluation differ and why? What are the requirements with respect to functionality of tools that may support collection of data from remote testing, and how do those requirements map on to currently available tools such as OpenVULab (used in this study)? Future studies that address these or other related questions may lend insight into how to maximize the ecological validity of accessibility evaluation of e-learning infrastructure by determining how to best support testing conducted in environments most closely resembling the actual working environments of students.

Finally, it is important to consider how to ensure that the findings of such studies may ultimately enhance the accessibility of e-learning in higher education. How can researchers effectively liaise with post-secondary practitioners, technologists, support services, staff developers, and senior managers to ensure that e-learning technologies in use, or being considered for use, are accessible for diverse student populations? An essential related area of study is therefore to develop effective frameworks (e.g., informed by the rainbow bridge metaphor for accessibility) that may be utilized on an institutional scale to affect change that will ultimately have a direct and positive impact on the e-learning experiences of all students. It is in this manner that we can begin to work practically towards designing for diversity of our student populations.

REFERENCES

Abdi, H. (2007). Bonferroni and Sidak corrections for multiple comparisons. In N. J. Salkind (Ed.), *Encyclopedia of measurement and statistics* (pp. 103-107). Thousand Oaks, CA: Sage.

Abrami, P. C., Bernard, R. M., Wade, A., Schmid, R. F., Borokhovski, E., Tamim, R., . . . Peretiatkowicz, A. (2006). A review of e-learning in Canada: A rough sketch of the evidence, gaps and promising directions. *Canadian Journal of Learning and Technology, 32*(3).

Advisory Committee for Online Learning. (2001). *The e-learning e-volution in colleges and universities*. Ottawa, ON: Council of Minister of Education.

AFB. (2008). Distance learning: How accessible are online educational tools. Retrieved from http://www.afb.org/section.aspx?FolderID=3&SectionID=3&TopicID=138&DocumentID=4492

Allen, I. E., & Seaman, J. (2010). Learning on demand: Online education in the United States, 2009. Retrieved from www.sloan-c.org/publications/survey/pdf/learningondemand.pdf

Allen, I. E., & Seaman, J. (2011). *Going the distance: Online education in the United States 2011*. Wellesley, MA: Babson Survey Research Group.

Alonso, F., Fuertes, J. L., González, A. L., & Martinez, L. (2010, April). *On the testability of WCAG 2.0 for beginners.* Paper presented at the 2010 International Cross Disciplinary Conference on Web Accessibility, Raleigh, NC.

Americans with Disabilities Act. (1990, 2008). ADA. Retrieved from http://www.ada.gov/pubs/ada.htm

Andreason, M. S., Villemann Nielsen, H., Ormholt Schroeder, S., & Stage, J. (2007). What happened to remote usability testing? An empirical study of three methods. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 1405-1414). New York, NY: ACM.

AODA Integrated Accessibility Standards. (2011). Retrieved from the Service Canada website: http://www.e-laws.gov.on.ca/html/source/regs/english/2011/elaws_src_regs_r11191_e.htm#BK15

Armstrong, H. L. (2009). Advanced IT education for the vision impaired via e-learning. *Journal of Information Technology Education, 8*, 243-256.

Arroway, P., Davenport, E., & Xu, G. (2010). EDUCAUSE core data service: Fiscal year 2009 summary report. Retrieved from http://www.educause.edu/ir/library/pdf/PUB8007.pdf

Asuncion, J. V., Fichten, C. S., & Barile, M. (2007). Which forms of e-learning are accessible to Canadian postsecondary students with disabilities? *Communique, 7*(3), 36.

Australian Government. (2011). Accessibility - Web guide. Retrieved from http://webguide.gov.au/accessibility-usability/accessibility/

Baravalle, A., & Lanfranchi, V. (2003). Remote Web usability testing. *Behavior Research Methods, Instruments, & Computers, 35*(3), 364-368.

Bartek, V., & Cheatham, D. (2003a). Experience remote usability testing, Part I. Retrieved from http://www.ibm.com/developerworks/web/library/wa-rmusts1/

Bartek, V., & Cheatham, D. (2003b). Experience remote usability testing, Part II. Retrieved from http://www.ibm.com/developerworks/web/library/wa-rmusts2/

Bates, A. W. (2005). *Technology, e-learning and distance education* (2nd ed.). New York, NY: Routledge.

Beiser, M., Erickson, D., Fleming, J. A. E., & Iacono, W. G. (1993). Establishing the onset of psychotic illness. *The American journal of psychiatry, 150*(9), 1349-1354.

Bernard, R. M., Abrami, P. C., Lou, Y., & Borokhovski, E. (2004). A methodological morass? How we can improve quantitative research in distance education. *Distance Education, 25*(2), 175-198.

Bernard, R. M., Abrami, P. C., Lou, Y., Borokhovski, E., Wade, A., Wozney, L., . . . Huang, B. (2004). How does distance education compare with classroom instruction? A meta-analysis of the empirical literature. *Review of Educational Research, 74*(3), 379-439.

Bishop, D., & Rhind, D. J. A. (2011). Barriers and enablers for visually impaired students at a UK higher education institution. *The British Journal of Visual Impairment, 29*(3), 177-195.

Blacklock, B., Benson, B., & Johnson, D. (2003). Needs assessment project: Exploring barriers and opportunities for college students with psychiatric disabilities. Retrieved from http://ds.umn.edu/outreach.html

Bland, M. (2000). Multiple significance tests and the Bonferroni correction *An introduction to medical statistics* (3rd ed.). New York, NY: Oxford Medical Publications.

Bochner, S. (1986). Observational methods. In W. J. Lonner & J. W. Berry (Eds.), *Field methods in cross-cultural research* (pp. 165-201). Beverly Hills, CA: Sage.

Bohman, P. R., & Anderson, S. (2005). *A conceptual framework for accessibility tools to benefit users with cognitive disabilities.* Paper presented at the W4A, Chiba, Japan.

Bolt, N., & Tulathimutte, T. (2010). *Remote Research.* Brooklyn, NY: Rosenfeld Media.

Boren, T., & Ramey, J. (2000). Thinking aloud: Reconciling theory and practice. *IEEE Transactions on Professional Communications, 43*(3), 261-278.

Brajnik, G. (2004). Comparing accessibility evaluation tools: A method for tool effectiveness. *Universal Access in the Information Society, 3*(3-4), 252-263.

Brajnik, G. (2009). Validity and reliability of web accessibility guidelines. In *Proceedings of the 11th International ACM SIGACCESS Conference on Computers and Accessibility* (pp. 131-138). New York, NY: ACM.

Brajnik, G., Yesilada, Y., & Harper, S. (2012). Is accessibility conformance an elusive property? A study of validity and reliability of WCAG 2.0. *ACM Transactions on Accessible Computing (TACCESS), 4*(2), 8:1-8:28.

Brown, R. S. (2008). Special education / special needs information in the TDSB, 2005-6. Retrieved from http://www.tdsb.on.ca/wwwdocuments/about_us/external_research_application/docs/SpecialEducationSpecialNeedsInformationInTheTDSB2005-6-%20EmailVersion.pdf

Burgstahler, S., Corrigan, B., & McCarter, J. (2004). Making distance learning courses accessible to students and instructors with disabilities: A case study. *Internet and Higher Education, 7*(3), 233-246.

Bury, S., & Oud, J. (2004). Usability testing of an online information literacy tutorial. *Reference Services Review, 33*(1), 54-65.

Cameron, A. C., & Trivedi, P. K. (1998). *Regression analysis of count data.* New York, NY: Cambridge University Press.

Campbell, D. T., & Stanley, J. C. (1966). *Experimental and quasi-experimental designs for research.* Chicago, IL: Rand McNally.

Canadian Council on Learning. (2009). *State of e-learning in Canada.* Ottawa, ON: Canadian Council on Learning.

CAST. (2011). About UDL. Retrieved from http://www.udlcenter.org/aboutudl

Center for Universal Design. (1997). The principles of universal design. Retrieved from http://www.ncsu.edu/www/ncsu/design/sod5/cud/about_ud/udprinciplestext.htm

Clark, J. (2006, May 23). To hell with the WCAG2. *A List Apart, 217.* Retrieved from http://www.alistapart.com/articles/tohellwithwcag2

CMHA. (n.d.-a). Facts about mental illness. Retrieved from http://www.cmha.ca/media/fast-facts-about-mental-illness/

CMHA. (n.d.-b). Understanding mental illness. Retrieved from http://cmha.ca/bins/content_page.asp?cid=3&lang=1

Cobus, L., Dent, V. F., & Ondrusek, A. (2005). How twenty-eight users helped redesign an academic library Web site: A usability study. *Reference & User Services Quarterly, 44*(3), 232-246.

Comeaux, D., & Schmetzke, A. (2007). Web accessibility trends in university libraries and library schools. *Library Hi Tech, 25*(4), 457-477.

Cook, R. A., & Gladhart, M. A. (2002). A survey of online instructional issues and strategies for postsecondary students with learning disabilities. *Information Technology and Disabilities, 8*(1).

Coombs, N. (2010). *Making online teaching accessible: Inclusive course design for students with disabilities*. San Francisco, CA: Jossey-Bass.

Cooper, M., Sloan, D., Kelly, B., & Lewthwaite, S. (2012). A challenge to web accessibility metrics and guidelines: Putting people and processes first. In *Proceedings of the International Cross-Disciplinary Conference on Web Accessibility* (pp. 1-4). New York, NY: ACM.

D'Agostino, R. B., Sullivan, L. M., & Beiser, A. S. (2006). *Introductory applied statistics*. Belmont, CA: Duxbury.

Dardailler, D. (2009). WAI early days. Retrieved from http://www.w3.org/WAI/history

Denton, W., & Coysh, S. J. (2011). Usability testing of VuFind at an academic library. *Library Hi Tech, 29*(2), 301-319.

Dermody, K., & Majekodunmi, N. (2010). Online databases and the research experience for university students with print disabilities. *Library Hi Tech, 29*(1), 149-160.

Dray, S., & Siegel, D. (2004). Remote possibilities? International usability testing at a distance. *Interactions, 11*(2), 10-17.

DRC. (2004). The Web: Access and inclusion for disabled people. A formal investigation conducted by the Disability Rights Commission. Retrieved from www.city.ac.uk/__data/assets/pdf_file/0004/72670/DRC_Report.pdf

Elias, T. (2010). Universal Instructional Design principles for Moodle. *International Review of Research in Open and Distance Learning, 11*(2), 110-124.

Ellcessor, E. (2010). Bridging disability divides: A critical history of web content accessibility through 2001. *Information, Communication & Society, 13*(3), 289-308.

Elston, R. C., & Johnson, W. D. (2008). *Basic biostatistics for geneticists and epidemiologists: A practical approach*. Chichester, UK: Wiley.

Ericsson, K. A., & Simon, H. A. (1993). *Protocol analysis: Verbal reports as data*. Cambridge, MA: MIT Press.

Eudaly, J. (2003). A rising tide: Students with psychiatric disabilities seek services in record numbers. Retrieved from http://www.heath.gwu.edu/PDFs/Psychiatric%20Disabilities.pdf

Evett, L., & Brown, D. (2005). Text formats and web design for visually impaired and dyslexic readers - Clear text for all. *Interacting with Computers, 17*(4), 453-472.

Faulkner, L. (2003). Beyond the five-user assumption: Benefits of increased sample sizes in usability testing. *Behavior Research Methods, Instruments, & Computers, 35*(3), 379-383.

Fichten, C. S., Ferraro, V., Asuncion, J. V., Chwojka, C., Barile, M., Nguyen, M. N., . . . Wolforth, J. (2009). Disabilities and e-learning problems and solutions: An exploratory study. *Educational Technology & Society, 12*(4), 241-256.

Fichten, C. S., Jorgensen, S., Havel, A., & Barile, M. (2006). College students with disabilities: Their future and success. Final report presented to FQRSC (Fonds de recherche sur la société et la culture). Retrieved from http://www.adaptech.org/cfichten/abExecutiveSummaryCollegeStudentsWithDisabilities.pdf

Franklin, U. M. (2004). *The real world of technology*. Toronto, ON: House of Anansi Press Inc.

Friedman, M. G., & Bryen, D. N. (2007). Web accessibility design recommendations for people with cognitive disabilities. *Technology and Disability, 19*(4), 205-212.

Gladhart, M. A. (2010). Determining faculty needs for delivering accessible electronically delivered instruction in higher education. *Journal of Postsecondary Education and Disability, 22*(3), 185-196.

Glass, G., McGaw, B., & Smith, M. L. (1981). *Meta-analysis in social research*. Newbury Park, CA: Sage Publications.

Goggin, G., & Newell, C. (2003). *Digital disability: The social construction of disability in new media*. New York, NY: Rowman & Littlefield.

Grabinger, S. (2010). A framework for supporting postsecondary learners with psychiatric disabilities in online environments. *Electronic Journal of e-Learning, 8*(2), 101-110.

Grace, S., & Gravestock, P. (2009). *Inclusion and diversity: Meeting the needs of all students*. New York, NY: Routledge.

Guerico, J. M., & Dixon, M. R. (2011). The observer effect and its impact on staff behavior in an acquired brain injury neurobehavioral treatment setting. *Journal of Organizational Behavior Management, 31*(1), 43-54.

Halbach, T. (2010, February). *Towards cognitively accessible web pages*. Paper presented at the 3rd International Conference on Advances in Computer-Human Interactions, Saint Maarten, Netherlands, Antilles.

Hammontree, M., Weiler, P., & Nayak, N. (1994). Remote usability testing. *Interactions, 1*(3), 21-25.

Harper, K. A., & DeWaters, J. (2008). A quest for website accessibility in higher education institutions. *Internet and Higher Education, 11*(3), 160-164.

Hartson, H. R., Castillo, J. C., Kelso, J., & Neale, W. (1996). Remote evaluation: The network as an extension of the usability laboratory. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 228-235). New York, NY: ACM.

Harty, J. (2011). Automated usability tests can be valuable companions to in-person tests. *ACM Queue, 9*(1), 8.

Henry, S. L. (2007). Just ask: Integrating accessibility throughout design. Retrieved from http://www.uiaccess.com/JustAsk/index.html

Hertzum, M., Hansen, K. D., & Andersen, H. H. K. (2009). Scrutinising usability evaluation: Does thinking aloud affect behaviour and mental workload? *Behaviour & Information Technology, 28*(2), 165-181.

Higher Education Research Institute. (2011). College students with "hidden" disabilities. Retrieved from http://www.heri.ucla.edu/publications-main.php

Houck-Whitaker, J. (2005). Remote testing versus lab testing. Retrieved from http://www.boltpeters.com/articles/versus.html

Hox, J. J. (2010). *Multilevel analysis: Techniques and applications*. New York, NY: Routledge.

Huang, W.-H. D., & Nakazawa, K. (2010). An empirical analysis on how learners interact in wiki in a graduate level online course. *Interactive Learning Environments, 18*(3), 233-244.

Hubka, D., & Killean, E. (1999). Working towards a coordinated national approach to services, accommodations and policies for post-secondary students with disabilities. Retrieved from http://www.neads.ca/en/about/projects/nasp/

Hughes, B., & Paterson, K. (1997). The social model of disability and the disappearing body: Towards a sociology of impairment. *Disability & Society, 12*(3), 325-340.

Hutchinson, N. L. (2010). *Inclusion of exceptional learners in Canadian schools* (3rd ed.). Toronto, ON: Pearson.

Hyde, M., Punch, R., Power, D., Hartley, J., Neale, J., & Brennan, L. (2009). The experiences of deaf and hard of hearing students at a Queensland University: 1985–2005. *Higher Education Research & Development, 28*(1), 85-98.

IMS Global Learning Consortium. (2004). IMS access for all metadata overview. Retrieved from http://www.imsglobal.org/accessibility/accmdv1p0/imsaccmd_oviewv1p0.html

ISO 9241. (1998). *Ergonomics requirements for office with visual display terminals (VDTs)*. Geneva: International Organization for Standardization.

Joint Department of Justice and Department of Education. (2011). Frequently asked questions about the June 29, 2010, Dear Colleagues letter. Retrieved from http://www2.ed.gov/about/offices/list/ocr/docs/dcl-ebook-faq-201105.html

Jung, S., Herlocker, J. L., Webster, J., Mellinger, M., & Frumkin, J. (2008). LibraryFind: System design and usability testing of academic metasearch system. *Journal of the American Society for Information Science and Technology, 59*(3), 375-389.

Kapsi, M., Vlachogiannis, E., Darzentas, J., & Spyrou, T. (2009a). *A preliminary feedback for the WCAG 2.0-WCAG 1.0 Vs WCAG 2.0 evaluation study.* Paper presented at the 2nd Annual Conference on Pervasive Technologies Related to Assistive Environments (PETRA), Confu, Greece.

Kapsi, M., Vlachogiannis, E., Darzentas, J., & Spyrou, T. (2009b). The usability of Web accessibility guidelines: An approach for evaluation. In *Proceedings of the 5th International Conference on Universal Access in Human Computer Interaction* (pp. 716-724). Berlin, Germany: Springer-Verlag.

Kelly, B., Phipps, L., & Swift, E. (2004). Developing a holistic approach for e-learning accessibility. *Canadian Journal of Learning and Technology, 30*(3).

Kelly, B., Sloan, D., Phipps, L., Petrie, H., & Hamilton, F. (2005). Forcing standardization or accommodating diversity? A framework for applying the WCAG in the real world. In *Proceedings of the 2005 International Cross-Disciplinary Workshop on Web Accessibility* (pp. 46-54). New York, NY: ACM.

Kennedy, H., Evans, S., & Thomas, S. (2011). Can the Web be made accessible for people with intellectual disabilities? *The Information Society, 27*(1), 29-39.

Kessler, R. C., Amminger, G. P., Aguilar-Gaxiola, S., Alonso, J., Lee, S., & Ustun, T. B. (2007). Age of onset of mental disorders: A review of recent literature. *Current Opinion in Psychiatry, 20*(4), 359-364.

Kim, P., Hong, J.-S., Bonk, C. J., & Lim, G. (2011). Effects of group reflection variations in project-based learning integrated in a Web 2.0 learning space. *Interactive Learning Environments, 19*(4), 333-349.

Kraus, F. S. H. (2003). Methodology for remote usability activities. *IBM Systems Journal, 42*(4), 584-593.

Kuzon, W. M., Urbanchek, M. G., & McCabe, S. (1996). The seven deadly sins of statistical analysis. *Annals of Plastic Surgery, 37*(3), 265-272.

Laff, M., & Rissenberg, M. (2007). *Cognitive ability measures for accessible web content.* Paper presented at the 4th International Conference on Universal Access in Human Computer Interaction (HCI), Beijing, China.

Lang, H. G. (2002). Higher education for deaf students: Research priorities in the new millennium. *Journal of Deaf Studies and Deaf Education, 7*(4), 267-280.

Lang, H. G., McKee, B. G., & Conner, K. (1993). Characteristics of effective teachers: A descriptive study of the perceptions of faculty and deaf college students. *American Annals of the Deaf, 138*(3), 252-252.

LDAC. (2002). Official definition of learning disabilities. Retrieved from http://www.ldac-acta.ca/en/learn-more/ld-defined.html

LDAC. (n.d.). Prevalence of learning disabilities. Retrieved from http://www.ldac-acta.ca/en/learn-more/ld-basics/prevalence-of-lds.html

Lewin-Jones, J., & Hodgson, J. (2004). Differentiation strategies relating to the inclusion of a student with a severe visual impairment in higher education (modern foreign languages). *The British Journal of Visual Impairment, 22*(1), 32-36.

Lewis, J. R. (2006). Usability testing. In G. Salvendy (Ed.), *Handbook of Human Factors and Ergonomics* (pp. 1275-1316). Hoboken, NJ: John Wiley.

Lovett, B. J. (2010). Extended time testing accommodations for students with disabilities: Answers to five fundamental questions. *Review of Educational Research, 80*(4), 611-638.

Lowe, B., & Laffey, D. (2011). Is Twitter for the birds? Using Twitter to enhance student learning in a marketing course. *Journal of Marketing Education, 33*(2), 183-192.

Marks, D. (1997). Models of disability. *Disability and Rehabilitation, 19*(3), 85-91.

Marschark, M., Sapere, P., Convertino, C., & Seewagen, R. (2005). Access to postsecondary education through sign language interpreting. *Journal of Deaf Studies and Deaf Education, 10*(1), 38-50.

Masling, J., & Stern, G. (1969). Effect of the observer in the classroom. *Journal of Educational Psychology, 60*(5), 351-354.

McCarthy, J. E., & Swierenga, S. J. (2010). What we know about dyslexia and web accessibility: A research review. *Universal Access in the Information Society, 9*(2), 147-152.

McFadden, E., Hager, D. R., Elie, C. J., & Blackwell, J. M. (2002). Remote usability evaluation: Overview and case studies. *International Journal of Human-Computer Interaction, 14*(3&4), 489-502.

Monahan, P. (2010). Building a more engaged university: Strategic directions for York University 2010-2020. [White paper]. Retrieved from http://vpacademic.yorku.ca/whitepaper/docs/White_Paper_Overview_April_15.pdf

Mowbray, C. T., Bybee, D., & Collins, M. E. (2001). Follow-up client satisfaction in a supported education program. *Psychiatric Rehabilitation Journal, 24*(3), 237-247.

Napier, J., & Barker, R. (2004). Acessing unviersity education: Perceptions, preferences, and expectations for interpretation by Deaf students. *Journal of Deaf Studies and Deaf Education, 9*(2), 228-238.

NEADS. (2005). Inclusion of students with disabilities in college and university-sponsored activities. Retrieved from http://www.neads.ca/en/about/projects/inclusion/

Newell, C. (2008). Distance education: Enabling and disabling? In T. Evans, M. Haughey & D. Murphy (Eds.), *International handbook of distance education* (pp. 87-106). Bingley, UK: Emerald Group Publishing.

Nicolle, C., & Paulson, D. (2004). Making the Internet accessible by people with cognitive and communication impairments. *Universal Access in the Information Society, 3*(1), 48-56.

Nielsen, J. (1993). *Usability engineering*. Boston, MA: AP Professional.

Nielsen, J. (2000). Why you only need to test with 5 users. Retrieved from http://www.useit.com/alertbox/20000319.html

Nielsen, J., & Landauer, T. K. (1993). A mathematical model of the finding of usability problems. In *Proceedings of the INTERACT '93 and CHI '93 Conference on Human Factors in Computing Systems* (pp. 206-213). New York, NY: ACM.

Offer, D., & Spiro, R. P. (1987). The disturbed adolescent goes to college. *Journal of American College Health, 35*(5), 209-214.

Office of the Premier Dalton McGuinty. (2010, March 8). Speech from the throne. [Transcript]. Retrieved from http://www.premier.gov.on.ca/news/event.php?ItemID=11282&Lang=En

Oliver, M. (1996). *Understanding disability: From theory to practice*. Houndmills: Macmillan.

Olmsted-Hawala, E. L., Murphy, E. D., Hawala, S., & Ashenfelter, K. T. (2010a). Think-aloud protocols: A comparison of three think-aloud protocols for use in testing data-dissemination web sites for usability. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 2381-2390). New York, NY: ACM.

Olmsted-Hawala, E. L., Murphy, E. D., Hawala, S., & Ashenfelter, K. T. (2010b). Think-aloud protocols: analyzing three different think-aloud protocols with counts of verbalized frustrations in a usability study of an information-rich Web site. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 60-66). New York, NY: ACM.

Ontario Human Rights Commission. (n.d.). The opportunity to succeed: Achieving barrier-free education for students with disabilities. Retrieved from http://www.ohrc.on.ca/en/resources/discussion_consultation/ConsultEduDisablty2?page=ConsultEduDisablty2-Contents.html

Owen Hutchison, J. S., Atkinson, K., & Orpwood, J. (1998). *Breaking down barriers: Access to further and higher education for visually impaired students.* Cheltenham: Stanley Thornes Ltd.

Owston, R. (2010). E-learning business case for York University. Retrieved from http://irlt.yorku.ca/reports/E-learningcasefinalversion.pdf

Palloff, R. M., & Pratt, K. (1999). *Building learning communities in cyberspace: Effective strategies for the online classroom.* San Francisco, CA: Jossey-Bass.

Persson, A.-C., Langh, M., & Nilsson, J. (2010). Usability testing and redesign of library Web pages at Lund University, Faculty of Engineering: A case study applying a two-phase, systematic quality approach. *Information Research, 15*(2).

Petrie, H., Hamilton, F., King, N., & Pavan, P. (2006). Remote usability evaluations with disabled people. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 1133-1141). New York, NY: ACM.

Pledge, D. S., Lapan, R. T., Heppner, P. P., Kivlighan, D., & Roehlke, H. J. (1998). Stability and severity of presenting problems at a university counseling center: A 6-year analysis. *Professional Psychology: Research and Practice, 29*(4), 386-389.

Power, C., Petrie, H., & Mitchell, R. (2009). A framework for remote user evaluation of accessibility and usability of websites. In C. Stephanidis (Ed.), *Lecture Notes in Computer Science: Vol. 5614. Universal Access in Human Computer Interaction. Addressing Diversity* (pp. 594-601). Berlin, Germany: Springer-Verlag.

Power, C., Petrie, H., Sakharov, V., & Swallow, D. (2010). Virtual learning environments: Another barrier to blended and e-learning. In *Proceedings of the 12th International Conference on Computers Helping People with Special Needs: Part I* (pp. 519-526). Berlin, Germany: Springer-Verlag.

Rainger, P. (2003). A dyslexic perspective on e-content accessibility. Retrieved from http://www.jisctechdis.ac.uk/techdis/resources/detail/learnersmatter/Jan03_Dyslexia

Rangin, H. (2009). How-to guide for creating accessible online learning content (CANNect Technical Report). Retrieved from http://sloanconsortium.org/cannect/projectone/index.php

Rehabilitation Act. (1998). Rehabilitation Act of 1973. Retrieved from http://www2.ed.gov/policy/speced/reg/narrative.html

Rehabilitation Act. (2000). Section 508 standards. Retrieved from http://www.section508.gov/index.cfm?fuseAction=stdsdoc

Reid, L. G., & Snow-Weaver, A. (2008). WCAG 2.0: A web accessibility standard for the evolving web. In *Proceedings of the 2008 International Cross-Discipinary Workshop on Web Accessibility* (pp. 109-115). New York, NY: ACM.

Ribera, M., Porras, M., Boldu, M., Termens, M., Sule, A., & Paris, P. (2009). Web Content Accessibility Guidelines 2.0: A further step towards accessible digital information. *Program: Electronic Library and Information Systems, 43*(4), 392-406.

RNIB. (2008). Accessibility, usability, and inclusive design - Software Access Centre, RNIB. Retrieved from http://www.rnib.org.uk/professionals/softwareandtechnology/softwareaccesscentre/background/Pages/access_usability.aspx

Rose, D. H., & Meyer, A. (2002). *Teaching Every Student in the Digital Age: Universal Design for Learning.* Alexandria, VA: ASCD.

Rubin, J., & Chisnell, D. (2008). *Handbook of usability testing.* Indianapolis, IN: Wiley.

Samarawickrema, G., & Stacey, E. (2007). Adopting web-based learning and teaching: A case study in higher education. *Distance Education, 28*(3), 313-333.

Sapp, W. (2009). Universal design: Online educational media for students with disabilities. *Journal of Visual Impairment & Blindness, 103*(8), 495-500.

Seale, J. K. (2006a). *E-learning and disability in higher education: Accessibility research and practice*. New York, NY: Routledge.

Seale, J. K. (2006b). The rainbow bridge metaphor as a tool for developing accessible e-learning practices in higher education. *Canadian Journal of Learning and Technology, 32*(2).

Seale, J. K., & Cooper, M. (2010). E-Learning and accessibility: An exploration of the potential role of generic pedagogical tools. *Computers & Education, 54*(4), 1107-1116.

Seale, J. K., Draffan, E. A., & Wald, M. (2004). Exploring disabled learners' experiences of learning, LEXDIS final report to JISC. Retrieved from www.lexdis.ecs.soton.ac.uk/project/media/LEXDIS_ProjectReport_Dec08final.doc

Seeman, L. (2006a). Re: Formal objection to WCAG 2.0. [Electronic mailing list message]. Retrieved from http://lists.w3.org/Archives/Public/w3c-wai-gl/2006AprJun/0368.html

Seeman, L. (2006b). The techniques document is hard to follow. [Electronic mailing list message]. Retrieved from http://lists.w3.org/Archives/Public/w3c-wai-gl/2006AprJun/0184.html

Selvaraj, P. (2004). *Comparative study of synchronous remote and traditional in-lab usability evaluation methods.* Unpublished Master's thesis, Virginia Polytechnic Institute and State University. Retrieved from http://scholar.lib.vt.edu/theses/available/etd-05192004-122952/unrestricted/Thesis_Prakaash_Selvaraj.pdf

Sharpe, M. N., Bruininks, B. D., Blacklock, B., Benson, B., & Johnson, D. A. (2004). The emergence of psychiatric disabilities in postsecondary education. *NCSET Issue Briefs, 3*(1).

Shaughnessy, J. J., Zechmeister, E. B., & Zechmeister, J. S. (2000). *Research methods in psychology*. New York, NY: McGraw Hill.

Sloan, D., Heath, A., Hamilton, F., Kelly, B., Petrie, H., & Phipps, L. (2006). Contextual web accessibility: Maximizing the benefit of accessibility guidelines. In *Proceedings of the 2006 International Cross-Disciplinary Workshop on Web Accessibility* (pp. 121-131). New York, NY: ACM.

Sloan, D., & Kelly, B. (2008). *Reflections on the development of a holistic approach to web accessibility.* Paper presented at the 2008 Accessible Design in the Digital World Conference (ADDW), York, UK.

Smith, A., Rainie, L., & Zickuhr, K. (2011). College students and technology. Retrieved from http://www.pewinternet.org/Reports/2011/College-students-and-technology.aspx

Smith, S. D., Saloway, G., & Caruso, J. B. (2010). ECAR study of undergraduate students and information technology. Retrieved from http://www.educause.edu/Resources/TheECARStudyofUndergraduateStu/187215

Spano, R. (2006). Observer behavior as a potential source of reactivity: Describing and quantifying observer effects in a large-scale observational study of police. *Sociological Methods & Research, 34*(4), 521-553.

Special Educational Needs and Disability Act 2001. (2001, c. 10). Retrieved from the National Archives website of the UK: http://www.legislation.gov.uk/ukpga/2001/10 website:

Spool, J., & Schroeder, W. (2001). Testing web sites: Five users is nowhere near enough. In *CHI '01 Extended Abstracts on Human Factors in Computing Systems* (pp. 285-286). New York, NY: ACM Press.

Strangman, N., Hall, T., & Meyer, A. (2003). *Graphic organizers with UDL*. Wakefield, MA: National Center on Accessing the General Curriculum.

Tamim, R., Bernard, R. M., Borokhovski, E., Abrami, P. C., & Schmid, R. F. (2011). What forty years of research says about the impact of technology on learning: A second-order meta-analysis and validation study. *Review of Educational Research, 81*(1), 4-28.

Tamler, H. M. (2001). High-tech vs high-touch: Some limits of automation in diagnostic usability testing. Retrieved from http://www.htamler.com/papers/techtouch

Tang, J. C., Liu, S. B., Muller, M., Lin, J., & Drews, C. (2006). Unobtrusive but invasive: Using screen recording to collect field data on computer-mediated interaction. In *Proceedings of the 2006 20th Anniversary Conference on Computer Supported Cooperative Work* (pp. 479-482). New York, NY: ACM.

Termens, M., Ribera, M., Porras, M., Boldú, M., Sulé, A., & Paris, P. (2009). Web content accessibility guidelines: From 1.0 to 2.0. In *Proceedings of the 18th International Conference on the World Wide Web* (pp. 1171-1172). New York, NY: ACM.

Terzi, L. (2004). The social model of disability: A philosophical critique. *Journal of Applied Philosophy, 21*(2), 141-157.

Thatcher, J., Burks, M. R., Heilmann, C., Henry, S. L., Kirkpatrick, A., Lauke, P. H., . . . Waddell, C. D. (2006). Web accessibility: Web standards and regulatory compliance. Retrieved from http://library.books24x7.com.ezproxy.library.yorku.ca/assetviewer.aspx?bookid=14675&chunkid=335181448&rowid=1352&noteMenuToggle=0&leftMenuState=1

Thompson, T., Burgstahler, S., & Moore, E. J. (2010). Web accessibility: A longitudinal study of college and university home pages in the northwestern United States. *Disability & Rehabilitation: Assistive Technology, 5*(2), 108-114.

Thomsett-Scott, B. (2004). Yeah, I found it! Performing Web site usability testing to ensure that off-campus students can find the information they need. *Journal of Library Administration, 41*(3-4), 471-483.

Titchkosky, T. (2001). Disability: A rose by any other name? "People-first" language in Canadian society. *The Canadian Review of Sociology, 38*(2), 125-140.

Trace Center. (2007). About the Trace Center: History. Retrieved from http://trace.wisc.edu/about/

Treasury Board of Canada Secretariat. (2011). Standard on Web accessibility. Retrieved from http://www.tbs-sct.gc.ca/pol/doc-eng.aspx?section=text&id=23601

U.S. Department of Justice. (2008). Accessibility of state and local government websites to people with disabilities. Retrieved from http://www.ada.gov/websites2.htm

U.S. Department of Justice. (2010a). Letter of Resolution, D.J. No. 202-48-213 Princeton University. Retrieved from http://www.ada.gov/princeton.htm

U.S. Department of Justice. (2010b). Settlement agreement between the United States of America, the National Federation of the Blind, Inc. ("NFB"), and the American Council of the Blind ("ACB"), and the Arizona Board of Regents ("ABOR"), for and on behalf of Arizona State University ("ASU"). Retrieved from http://www.ada.gov/arizona_state_university.htm

U.S. Government Accountability Office. (2009). Higher education and disability: Education needs a coordinated approach to improve its assistance to schools in supporting students (No. GAO-10-33). Retrieved from http://161.203.16.70/products/GAO-10-33

Upton, G., & Cook, I. (2008a). Poisson distribution. Oxford University Press, A Dictionary of Statistics. Retrieved from http://www.oxfordreference.com.ezproxy.library.yorku.ca/view/10.1093/acref/9780199541454.001.0001/acref-9780199541454-e-1260

Upton, G., & Cook, I. (2008b). Regression.Oxford University Press, A Dictionary of Statistics. Retrieved from http://www.oxfordreference.com.ezproxy.library.yorku.ca/view/10.1093/acref/9780199541454.001.0001/acref-9780199541454-e-1378#.

Ustun, T. B., Chatterji, S., Bickenbach, J., Kostanjsek, N., & Schneider, M. (2003). The international classification for functioning, disability and health: A new tool for understanding disability and health. *Disability and Rehabilitation, 25*(11-12), 565-571.

Vanderheiden, G. (1995). Design of HTML (Mosaic) pages to increase their accessibility to users with disabilities: Strategies for today and tomorrow. Retrieved from http://trace.wisc.edu/archive/html_guidelines/version1.html

Vanderheiden, G. (2000). Fundamental principles and priority setting for universal usability. In *Proceedings of the 2000 Conference on Universal Usability* (pp. 32-38). New York, NY: ACM.

Vanderheiden, G. (2006). Cognitive and learning disability matrix. [Electronic mailing list message attachment]. Retrieved from http://lists.w3.org/Archives/Public/w3c-wai-gl/2006OctDec/att-0159/A-Cognitive_or_Learning_Disability.html

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

W3C. (1999). Web Content Accessibility Guidelines 1.0: W3C Recommendation 5-May-1999. Retrieved from http://www.w3.org/TR/WCAG10/

W3C. (2001). Web content accessibility guidelines 2.0: W3C working draft 25 January 2001. Retrieved from http://www.w3.org/TR/2001/WD-WCAG20-20010125/

W3C. (2005). Preliminary review of Web sites for accessibility. Retrieved from http://www.w3.org/WAI/eval/preliminary.html

W3C. (2008). Web Content Accessibility Guidelines 2.0: W3C Recommendation 11 December 2008. Retrieved from http://www.w3.org/TR/WCAG20/

W3C. (2009a). Facts about W3C. Retrieved from http://www.w3.org/Consortium/facts

W3C. (2009b). W3C mission. Retrieved from http://www.w3.org/Consortium/mission

WAI. (1997). WAI markup guidelines working group charter. Retrieved from http://www.w3.org/WAI/GL/charter.html

Wakefield, D. (2000). Interview on 508 regulations and the web. Retrieved from http://easi.cc/media/trans/tn508wld.htm

Waterfield, J., & West, B. (2002). *SENDA compliance in higher education: An audit and guidance tool for accessible practice within the framework of teaching and learning*: South West Academic Network for Disability Support (SWANDS), University of Plymouth.

WebAIM. (2009). Quick reference: Testing Web content for accessibility. Retrieved from http://webaim.org/resources/evalquickref/#human

WebAIM. (n.d.). WebAIM: Articles. Retrieved from http://webaim.org/articles/

Weiner, E., & Weiner, J. (1996). Concerns and needs of university students with psychiatric disabilities. *Journal of Postsecondary Education and Disability, 12*(1), 2-9.

White, B., & Larusson, J. A. (2010). Strategic directives for LMS planning. *Research Bulletin* (Vol. 19). Boulder, CO.

WHO. (2011). Disabilities. Retrieved from http://www.who.int/topics/disabilities/en/

Wilkoff, W., & Abed, L. (1994). *Practicing universal design: An interpretation of the ADA*. New York, NY: Van Nostrand Reinhold.

Williams, R. G., Klamen, D. A., & McGaghie, W. C. (2003). Cognitive, social, and environmental sources of bias in clinical performance settings. *Teaching and Learning in Medicine, 15*(4), 270-292.

Wilson, S. (1977). The use of ethnographic techniques in educational research. *Review of Educational Research, 47*(1), 245-265.

Winzer, M. (2008). *Children with exceptionalities in Canadian classrooms* (6th ed.). Toronto, ON: Pearson.

Wolanin, T. R., & Steele, P. E. (2004). Higher education opportunities for students with disabilities: A primer for policymakers. Retrieved from http://www.eric.ed.gov/ERICWebPortal/detail?accno=ED485430

Woodfine, B. P., Nunes, M. B., & Wright, D. J. (2008). Text-based synchronous e-learning and dyslexia: Not necessarily the perfect match! *Computers & Education, 50*(3), 703-717.

Woolrych, A., & Cockton, G. (2001). Why and when five test users aren't enough. In J. Vanderdonckt, A. Blandford & A. Derycke (Eds.), *Proceedings of IHM-HCI 2001 Conference* (Vol. 2, pp. 105-108). Toulouse, France: Cépadèus.

Yamamoto, K., & Suzuki, N. (2012). Effect of an observer's presence on facial behavior during dyadic communication. *Perceptual and Motor Skills, 114*(3), 949-963.

York University. (n.d.). How to Register | Mental Health Disability Services. Retrieved from http://mhds.info.yorku.ca/registration/

APPENDIX A: Sample Recruitment Flyer

PARTICIPANTS NEEDED FOR RESEARCH IN E-LEARNING ACCESSIBILITY

The purpose of this research is to learn how accessible or inaccessible e-learning is from the perspective of students. We are also interested in developing methods for students to participate in e-learning accessibility testing.

As a participant in this study, you will be asked to access a sample online course and to work though small learning modules in the course, while providing feedback about how easy or difficult it is to work with the webpages that you encounter within the course. During part of the session, your workstation (you and the computer that you will work from) will be videotaped in order to learn from your interactions with the computer. You will also be interviewed at the end of the session to learn more about your thoughts on the sample online course and your experience in participating in the testing session.

Your participation would involve a session of approximately 1 hour. Compensation in the form of a York University Bookstore gift card will be provided for participation.

To be eligible to participate in this study, you must be able to attend a classroom on campus at York University. You do not need to have advanced computer skills or any other special expertise. We are interested in participants with a wide range of learning styles and needs, and would like to recruit students who identify themselves as persons with learning disabilities as well as students who do not identify as persons with disabilities.

For more information about this study, or to volunteer for this study, please contact:

Kari Kumar
Graduate Student, Faculty of Education
Email: kari_kumar@edu.yorku.ca
or
Ron Owston
Professor, Faculty of Education
Email: ROwston@edu.yorku.ca
Tel: 416-736-5019

This study has received ethics approval from the Office of Research Ethics at York University.

E-Learning Accessibility Study
To learn more, contact:
kari_kumar@edu.yorku.ca

E-Learning Accessibility Study
To learn more, contact:
kari_kumar@edu.yorku.ca

E-Learning Accessibility Study
To learn more, contact:
kari_kumar@edu.yorku.ca

E-Learning Accessibility Study
To learn more, contact:
kari_kumar@edu.yorku.ca

E-Learning Accessibility Study
To learn more, contact:
kari_kumar@edu.yorku.ca

E-Learning Accessibility Study
To learn more, contact:
kari_kumar@edu.yorku.ca

E-Learning Accessibility Study
To learn more, contact:
kari_kumar@edu.yorku.ca

E-Learning Accessibility Study
To learn more, contact:
kari_kumar@edu.yorku.ca

E-Learning Accessibility Study
To learn more, contact:
kari_kumar@edu.yorku.ca

E-Learning Accessibility Study
To learn more, contact:
kari_kumar@edu.yorku.ca

APPENDIX B: Sample Screening Interview Questions

[Interviews took place by email or telephone.]

Thank you for your interest in our e-learning accessibility research study. In order to confirm your eligibility in the study, I'd like to ask you a few questions. Any information that you provide will remain confidential.

1. Do you identify as a person with a learning disability?

2. What types of challenges or frustrations, if any, do you commonly encounter when working on the computer, including when you access websites?

3. Do you require the use of assistive technologies when accessing websites on the computer?

4. Would it be possible for you to visit a classroom in the TEL building to participate in this study?

5. What days and times are you available to participate in this study during the period of *[insert date]* to *[insert date]*? We expect that participation would require about 1 hour of your time.

6. Do you have any questions about this study that you would like me to answer at this time?

Please reply to this email with your responses to the above questions or, if you prefer to discuss this over the telephone, please reply with a phone number that you can be reached at.

Sincerely,

Kari

Kari Kumar, PhD Candidate

Faculty of Education

York University

APPENDIX C: Scholarly Resources Module

SCHOLARLY RESOURCES MODULE HOMEPAGE

[The module was presented as a single Moodle page with steps A, B, and C as shown below.]

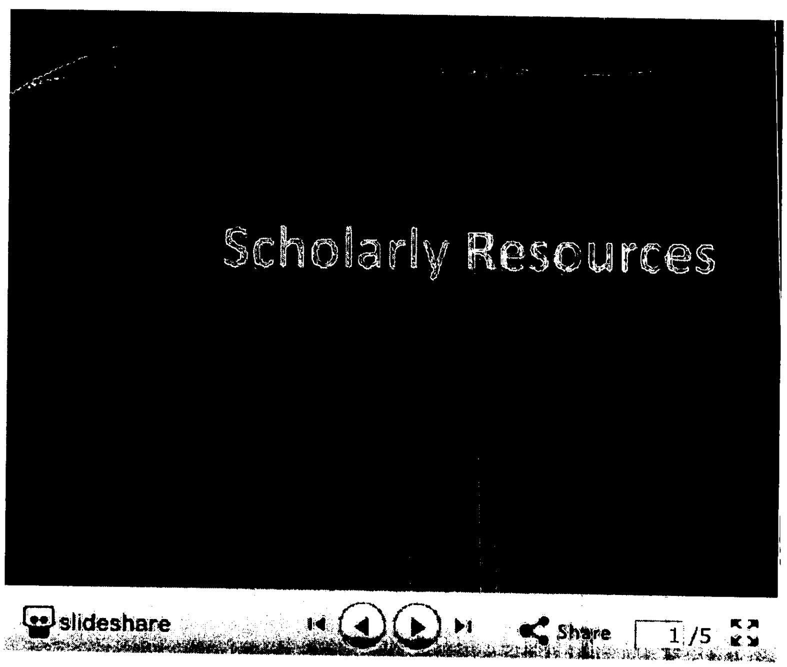
Upon successful completion of this module, you will have demonstrated the ability to:

- Locate a scholarly resource from an online library catalog; and to
- Access a current journal article.

<u>Please complete the steps below to complete this module:</u>

Step A: Learn about scholarly resources

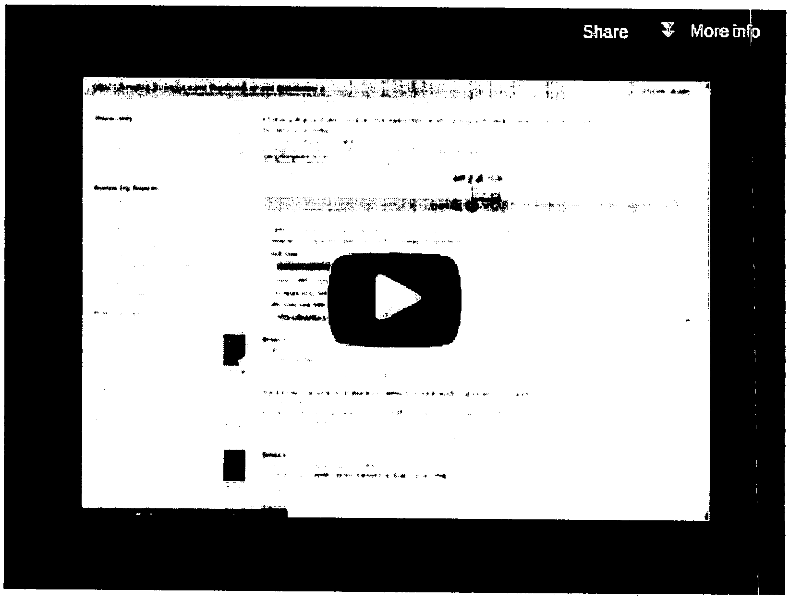
1. View a presentation to become acquainted with features of scholarly resources.



[Slides included within the presentation are shown in the next section of this appendix.]

Step B: Locate a journal article

1. View a Virginia Commonwealth University (VCU) library video demonstrating how to use the online catalog. *[The video is available from http://www.youtube.com/watch?feature=player_embedded&v=cbh8rgPo4yw]*



2. Use the **VCU Library Catalogue** to access the journal titled Academic Leadership. *[The catalogue is available at http://www.library.vcu.edu/]*
3. Skim the abstract of an article in the current issue of the journal for interest.

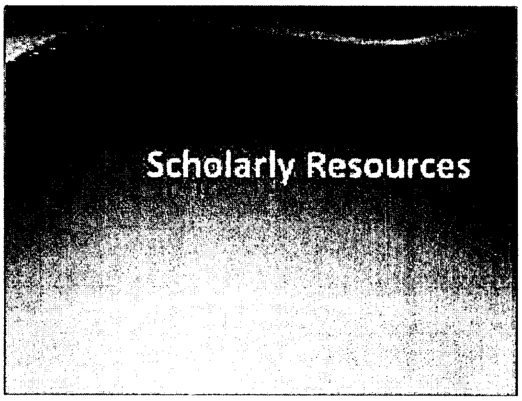
Step C: Discuss scholarly resources on a discussion forum

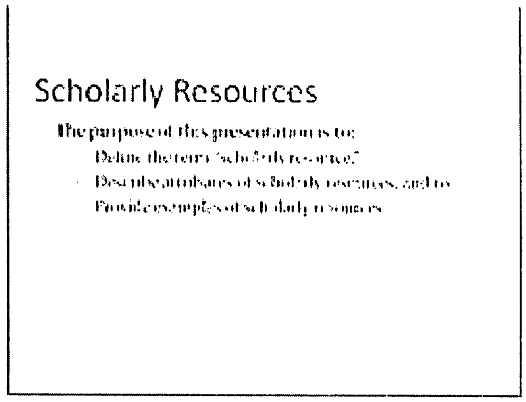
1. Think about the question "How easy or difficult was it to locate a scholarly source from the online VCU library catalog, and why?"
2. Post a brief comment about the question (approximately 2 sentences) on a designated discussion forum. **Posting this comment signifies completion of this module.

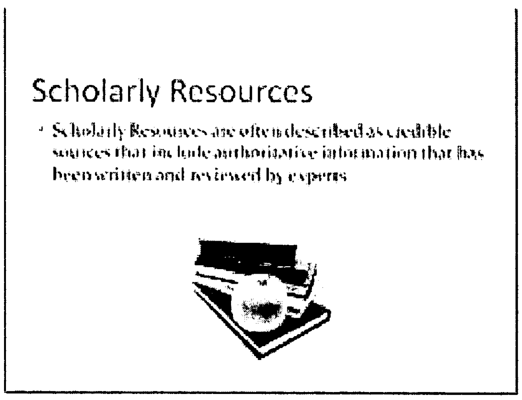
Scholarly Resources Discussion Forum

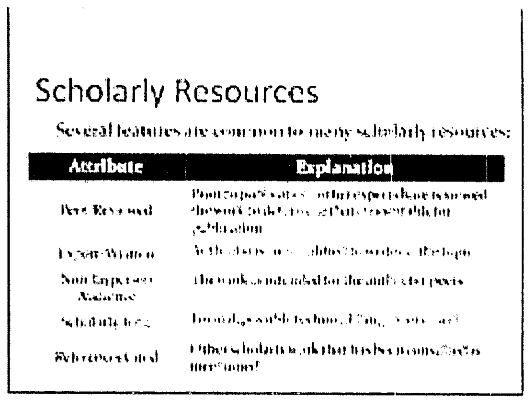
[The appearance of the discussion forum is shown in the next section of this appendix.]

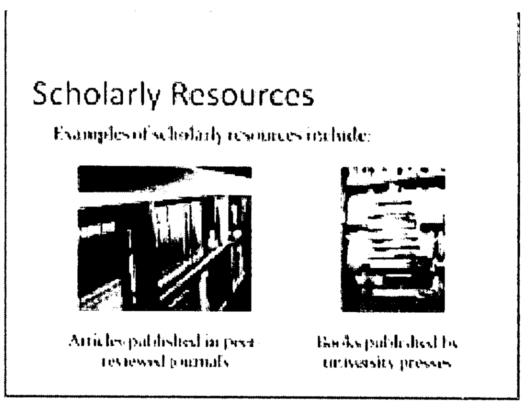
SCHOLARLY RESOURCES MODULE POWERPOINT SLIDES











SCHOLARLY RESOURCES DISCUSSION FORUM

Display replies in nested form

Retrieving sources

by Ken Hamer - Monday, 30 April 2012, 09:24 PM

Answer the question: How easy or difficult was it to locate a scholarly source from the online VCU library catalog, and why?

Edit | Delete | Reply

Re: Retrieving sources

by John Smith - Monday, 30 April 2012, 09:38 PM

It was pretty easy to find a journal from the library catalogue. I've done this lots of times for assignments

Show parent | Edit | Delete | Reply

Re: Retrieving sources

by Suzy Que - Monday, 30 April 2012, 09:39 PM

How did you do that John? I don't know what we were supposed to be looking for

Show parent | Edit | Delete | Reply

Re: Retrieving sources

by Barry Johnson - Monday, 30 April 2012, 09:40 PM

It was hard to find a scholarly source. There were too many options on the library website and I didn't know which link to click on

Show parent | Edit | Delete | Reply

[The discussion forum was pre-seeded with sample replies from fictitious students. Student participant replies were deleted so that all participants were presented with the same forum including the four posts shown above.]

APPENDIX D: Credible Resources module

CREDIBLE RESOURCES MODULE HOMEPAGE

[The module was presented as a single Moodle page with steps A, B, and C as shown below.]

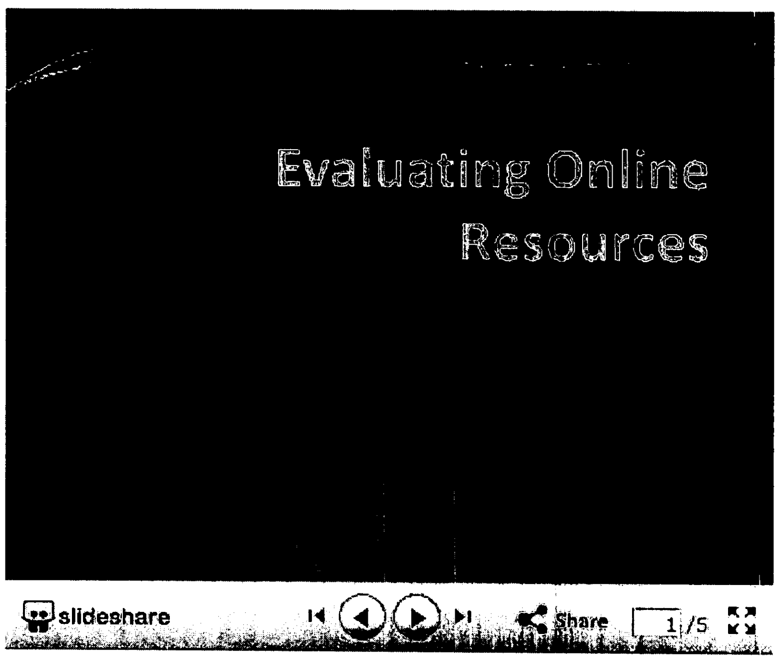
Upon successful completion of this module, you will have demonstrated the ability to:

- Determine whether an online resource is or is not credible; and to
- Articulate why a resource is or is not credible.

Please complete the steps below to complete this module:

Step A: Learn about how to evaluate online sources for credibility.

1. View a presentation to become acquainted with strategies to evaluate credibility of online sources.



[Slides included within the presentation are shown in the next section of this appendix.]

2. View a video that describes strategies to evaluate credibility of online sources. *[The video is available from http://www.youtube.com/watch?feature=player_embedded&v=9ig20cOGQYU]*

Step B: Visit websites and consider whether they contain credible sources of information (e.g., suitable for use in academic work).

1. Visit website 1, **VCU News Center article**, and examine the website for clues as to whether or not it may be a credible source of information. *[The website is http://www.news.vcu.edu/news/Potential_New_Therapeutic_Molecular_Target_to_Fight_Cancer]*
2. Visit website 2, **Yahoo Discussion Forum**, and examine the website for clues as to whether or not it may be a credible source of information. *[The forum is at http://answers.yahoo.com/question/index?qid=20080115181149AAKOAZ5]*

Step C: Discuss credible resources on a discussion forum

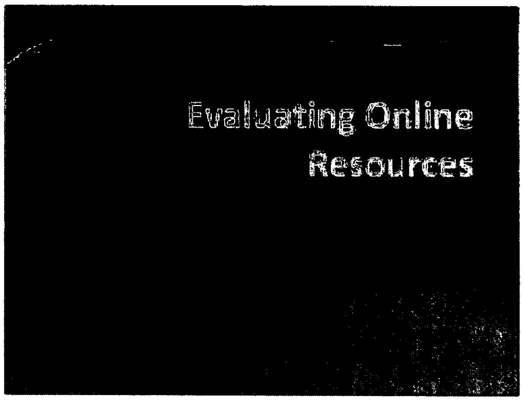
1. Think about the question "How easy or difficult was it for you to determine whether the websites you viewed are credible sources, and why?"

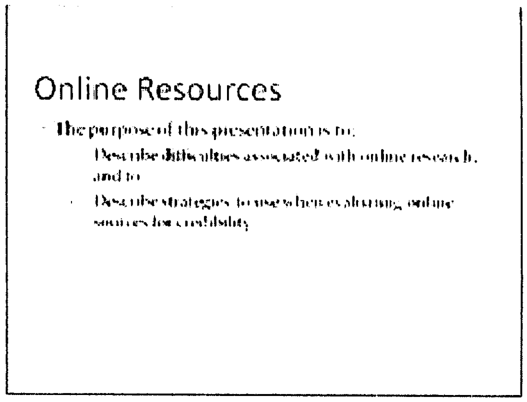
2. Post a brief comment about the question (approximately 2 sentences) on a designated discussion forum. **Posting this comment signifies completion of this module.

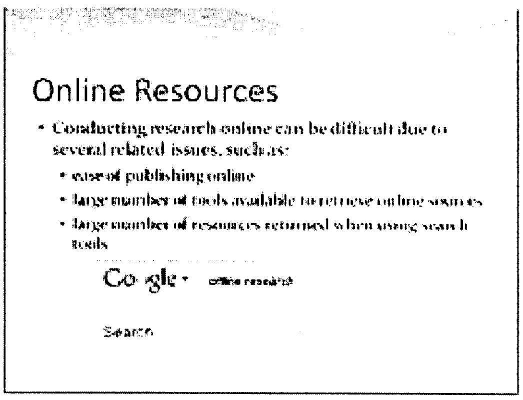
<u>**Credible Resources Discussion Forum**</u>

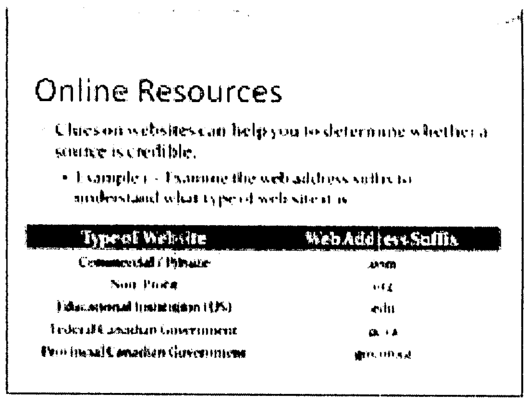
[The appearance of the discussion forum is shown in the next section of this appendix.]

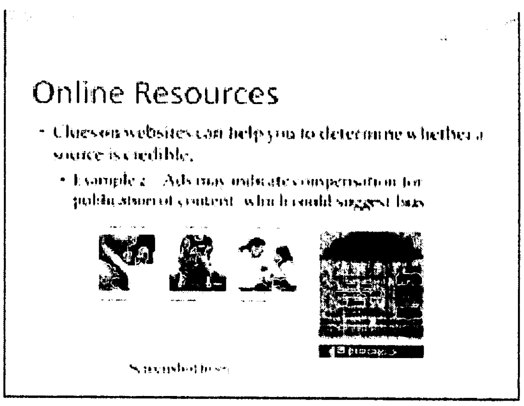
CREDIBLE RESOURCES MODULE POWERPOINT SLIDES











CREDIBLE RESOURCES DISCUSSION FORUM

Display replies in nested form

Online sources

by Kan Kumar - Monday, 30 April 2012, [illegible] PM

Answer the question How easy or difficult was it for you to determine whether the websites you viewed are credible sources, and why?

Edit | Delete | Reply

Re: Online sources

by John Seven - Monday, 30 April 2012, [illegible] PM

I wasn't able to tell which website was credible Are discussion forums ever credible?

Show parent | Edit | Delete | Reply

Re: Online sources

by Suzy Que - Monday, 30 April 2012, [illegible] PM

I know what you mean! I think that some discussion forums can be credible depending on who is posting there

Show parent | Edit | Delete | Reply

Re: Online sources

by Bailey Johnson - Monday, 30 April 2012, [illegible] PM

It was easy for me to compare the websites I found that one website had more formal language and that made it seem more credible to me

Show parent | Edit | Delete | Reply

[The discussion forum was pre-seeded with sample replies from fictitious students. Student participant replies were deleted so that all participants were presented with the same forum including the four posts shown above.]

APPENDIX E: Verbal Introduction Script

[Note that portions of this script have been taken verbatim, or nearly so, from Rubin and Chisnell (2008, pp. 156-157).]

Thank you for agreeing to take part in our e-learning accessibility research study. My name is Kari, and I'll be working with you today. Have you had an opportunity to read the consent form that I emailed to you? *[If not, present the student with a form to read. Collect a signed consent form before proceeding.]*

During the rest of our session, I'll be working from a script to ensure that my instructions to all students who participate in the study are the same.

I'm interested in learning how easy or difficult it is for students to work with technology as they complete online courses, as well as how comfortable students feel in participating in a study like this and giving feedback about their online learning experiences.

During this session, I'll ask you to access a sample online course in Moodle. This is not a real course – it is a mock course that was created for this study. The course is called "Introduction to Digital Literacy" and there are two small learning modules in the course. I will ask you to visit each of the learning modules separately, and attempt to access the materials and complete a task. As you do so, please try to do whatever you would normally do.

You are not being tested for your ability to complete the course modules – it's you who are helping me to evaluate the online course. I am interested in your thoughts as you work with the online course. Please try to think out loud while you are working. This will help me to understand what works or doesn't work within the course.

Here is how this session will be organized:

- There is a work station in the room with a computer for you to work with
- You'll work through each of the two sample course modules separately. First, you'll complete a short online questionnaire, and then you'll visit the course website. I'll help you to open up the questionnaire and ensure that the workstation is set up in a way that is comfortable and functional for you.
- I'm interested in learning whether or not students are able to participate in a session like this independently, or if it is better to have a moderator like myself present. For this reason:
    - During completion of one of the modules I will stay in the room and may prompt you to think out loud or even ask you a question or two. There will also be a video camera recording you as you work.
    - During completion of the other module, I will leave the room and will observe the session from the adjacent room by looking through the one-

way mirror. There will be no video camera used in this task. If you encounter a major problem such as loss of Internet connectivity, please signal to me by waving at the mirror so that I can visit you here.

- Completion of each module may take about 20 minutes, and you may take a break in between modules. It's possible that you may not be able to complete one or both modules, if you encounter inaccessible technology. This is okay, and I'll let you know when to stop your attempt.
- After finishing work with each module, you'll be asked to answer a few questions about your experience with it.
- After you have finished working with both modules, I will interview you about your experience and that will conclude our session.

Do you have any questions before we begin?

To begin, I'll first ask you to practice thinking out loud so that you become accustomed to this.

APPENDIX F: Think Aloud Introduction Script

I would like you to think out loud while you work with the online course today. Essentially, I would like you say anything that comes to your mind. This can feel unnatural at first, and so I'll first demonstrate this and then ask you to practice.

[Demonstration to be completed by test administrator while the student listens: The task is to locate a page to book a library study room. Begin on the testing computer from the York University homepage. Intentionally take an indirect route to finding the page, so that verbalizations will include false starts as well as explanations of problems encountered.]

The task that I'd like you to practice with is to start from the York University homepage and find a listing of the campus bookstore hours. Complete this task the way you normally would, and verbalize all of your thoughts along the way.

[The practice task was administered using OpenVULab so that students also had an opportunity to practice using the tool and completing pre- and post-session surveys.]

APPENDIX G: Exit Interview Protocol

Thank you for participating in this testing session today and working with the sample online course. To conclude our session today, I'd like to ask you a few questions to learn more about your experience participating in this study.

I'd first like to ask you for your general impressions about completing the modules.

[The module homepages were opened on the computer for reference.]

1. One of the modules that you worked on asked you to find a scholarly source - a journal article – and to then comment on that task on a discussion forum.
    - What was the most challenging aspect of completing this module? Can you tell me why this was challenging?
    - What was the least challenging (or easiest) aspect of completing this module? Can you tell me what made this aspect of the module the easiest?
2. One of the modules that you worked on asked you to view websites and to think about whether or not they were credible sources of information, and to then comment on that task on a discussion forum.
    - What was the most challenging aspect of completing this module? Can you tell me why this was challenging?
    - What was the least challenging (or easiest) aspect of completing this module? Can you tell me what made this aspect of the module the easiest?

You also participated in testing the online course in two arrangements: one where you were alone in the testing room, and another where I was there and a video camera was also present.

I'd first like to ask you about your general impressions about participating in the testing when you were alone in the room.

3. What was the most positive aspect of participating in the testing when you were alone in the room?
4. What was the least positive aspect of participating in that test?

I'd now like to ask you about your general impressions about participating in the testing when I was in the room with you and you were also videotaped.

5. What was the most positive aspect of participating in the testing when the video camera and I were there?
6. What was the least positive aspect of participating in that test?

7. If you were planning to participate again in this type of testing and were asked to choose which method to participate in, which one would you choose? Tell me why.

I'd like to ask you now about your thoughts on working on this sample course as an online course.

8. During the testing session today, were there instances where you felt that accessing this course online may have been easier for you to understand and complete the tasks than it may have been if this course was presented face-to-face?
9. During the testing session today, were there instances where you felt that accessing this course online may have been more difficult for you to understand and complete the tasks than it may have been if this course was presented face-to-face?

10. Is there anything else that you would like to tell me about regarding the sample online course, participating in the different testing arrangements, or e-learning in general?

Thank you for participating in this study. Your feedback has been very helpful today. [*Student will be presented with compensation.*]

APPENDIX H: Pre-Test Questionnaire

Part I: Questions common to tests involving both modules

1. What gender do you identify with?
    - o Male
    - o Female
    - o Other

2. What type of disability, if any, do you identify with? You may select more than one answer for this question.
    - o I do not have a disability
    - o I have a learning or cognitive disability
    - o I have a physical disability
    - o I have a mental health disability
    - o I have a hearing impairment
    - o I have a visual impairment

3. Have you taken an online course before?
    - o Yes
    - o No

4. How would you rate your overall computer skills?

   - o Very poor
   - o Poor
   - o Moderate
   - o Good
   - o Very good

5. How would you rate your ability to navigate through websites?

   - o Very poor
   - o Poor
   - o Moderate
   - o Good
   - o Very good

6. How would you rate your ability to read text on websites?

   - o Very poor
   - o Poor
   - o Moderate
   - o Good
   - o Very good

7. How would you rate your ability to understand written instructions?
    - o Very poor
    - o Poor
    - o Moderate
    - o Good
    - o Very good

8. How would you rate your ability to understand verbal instructions?
    - o Very poor
    - o Poor
    - o Moderate
    - o Good
    - o Very good

9. How would you rate your ability to understand visual instructions such as demonstrations?
    - o Very poor
    - o Poor
    - o Moderate
    - o Good
    - o Very good

10. How often have you used Moodle as a course website?

- o Never
- o Rarely
- o Occasionally
- o Frequently
- o Very frequently

11. How often have you viewed PowerPoint files for learning purposes?

- o Never
- o Rarely
- o Occasionally
- o Frequently
- o Very frequently

12. How often have you viewed videos for learning purposes?

- o Never
- o Rarely
- o Occasionally
- o Frequently
- o Very frequently

13. How would you rate your ability to express your thoughts in writing?

- Very poor
- Poor
- Moderate
- Good
- Very good

14. How often have you posted messages to online discussion forums for academic purposes?

- Never
- Rarely
- Occasionally
- Frequently
- Very frequently

Part IIa: Questions included only prior to Scholarly Resources module testing

15. How often have you used an online library catalogue system?

- o Never
- o Rarely
- o Occasionally
- o Frequently
- o Very frequently

16. How often have you accessed websites of academic journals?

- o Never
- o Rarely
- o Occasionally
- o Frequently
- o Very frequently

Part IIb: Questions included only prior to Credible Resources module testing

15. How often have you read descriptions of academic research projects?

    - Never
    - Rarely
    - Occasionally
    - Frequently
    - Very frequently

16. How often have you accessed online discussion forums that are not associated with a course webpage?

    - Never
    - Rarely
    - Occasionally
    - Frequently
    - Very frequently

APPENDIX I: Post-Test Questionnaire

1. Overall, how easy was it to work within Moodle while completing this module?

    - Very difficult
    - Difficult
    - Moderate
    - Easy
    - Very easy

2. How easy was it to read the instructions included within this module?

    - Very difficult
    - Difficult
    - Moderate
    - Easy
    - Very easy

3. How easy was it to understand the instructions included within this module?

    - Very difficult
    - Difficult
    - Moderate
    - Easy
    - Very easy

4. How easy was it to access the PowerPoint presentation within this module?
   - Very difficult
   - Difficult
   - Moderate
   - Easy
   - Very easy

5. How easy was it to understand the PowerPoint presentation within this module?
   - Very difficult
   - Difficult
   - Moderate
   - Easy
   - Very easy

6. How easy was it to access the video included within this module?
   - Very difficult
   - Difficult
   - Moderate
   - Easy
   - Very easy

7. How easy was it to understand the video included within this module?

    - Very difficult
    - Difficult
    - Moderate
    - Easy
    - Very easy

8. How challenging was the question included at the end of this module?

    - Very difficult
    - Difficult
    - Moderate
    - Easy
    - Very easy

9. How easy was it to post your answer to the question on the discussion forum in Moodle?

    - Very difficult
    - Difficult
    - Moderate
    - Easy
    - Very easy

10. How easy was it to work with websites outside of Moodle while completing this module?

- o Very difficult
- o Difficult
- o Moderate
- o Easy
- o Very easy

11. Overall, how easy was it to complete this module?

- o Very difficult
- o Difficult
- o Moderate
- o Easy
- o Very easy

12. Overall, how easy was it to participate in this testing session where you [*insert in brackets "worked in the room independently" or "worked in the presence of a researcher and video camera"*]?

- o Very difficult
- o Difficult
- o Moderate
- o Easy

- Very easy

13. Overall, how would you rate your comfort level with participating in this testing session where you [*insert in brackets "worked in the room independently" or "worked in the presence of a researcher and video camera"*]?
    - Very low
    - Low
    - Moderate
    - High
    - Very high

14. Please include any other comments that you wish to make about this testing session or the online course module below. [This question is optional]