Customization: The Path to a Better and More Accessible Web Experience

Ryan Fritz, Kim-Phuong L. Vu(✉), and Wayne E. Dick

California State University, Long Beach, CA 90840, USA
RYFRITZ2006@yahoo.com, Kim.Vu@CSULB.edu,
WAYNEEDICK@gmail.com

Abstract. The ever-increasing prevalence of electronic information online has provided a medium that allows for greater access to a wider range of users. It is important that user groups such as older adults, or individuals with visual impairments or disabilities, have an equal level of access to electronic information as do individuals without visual impairments or disabilities. This study examined whether allowing older adults the ability to customize various aspects of text would improve their reading performance and subjective usability ratings, while reducing the level of reported visual fatigue for online reading. Data from 16 older adult participants (age range 49 to 69) and 16 younger adult participants (age range 18 to 22) were analyzed. Participants were asked to read text passages using a big or small screen size, and answer reading comprehension questions under conditions where they were able to customize text or not able to customize text. Reading performance and ratings of usability and visual fatigue were obtained. Results showed an interaction between customization condition and age group. In the customized reading conditions, younger adults were more accurate than older adults in answering reading comprehension questions, but their performance did not differ significantly in the non-customized reading conditions. There was no effect of screen size on any of the dependent measures. Implications of these findings are discussed.

Keywords: Web-accessibility · Text-customization · Older adults

1 Introduction and Literature Review

Reading performance varies depending on whether an individual is reading from printed materials or from electronic sources. When reading from a computer screen, people typically take longer compared to when reading the same information from a printed source [1]. Despite the difference in reading speed, reading comprehension does not appear to differ across printed or electronic resources, suggesting that electronic content can be understood just as well as printed materials [2]. Although reading comprehension is similar, whether the electronic document can be useful depends on a variety of factors that impact usability and accessibility including, properties of the materials (font size, accessibility features), user characteristics (age, impairments, fatigue), and viewing conditions (mobile or desktop, lighting). In the following sections, an overview of the importance of usability and accessibility will be presented

S. Yamamoto and H. Mori (Eds.): HCII 2019, LNCS 11569, pp. 3–21, 2019.
https://doi.org/10.1007/978-3-030-22660-2_1

followed by a literature review on factors that influence user performance with reading electronic text on the web. Finally, results from an experiment conducted that examined user performance and preferences with electronic material that can be customized versus material that cannot be customized is presented.

1.1 Usability and Accessibility

Many aspects of our society such as education, banking, and healthcare have turned to the web as an additional medium through which organizations can interact with greater numbers of people. With this ever-increasing availability of information that can be obtained online, and its growing prevalence in our everyday lives, it is essential that individuals are able to access the information regardless of their mental or physical limitations. According to the United Nations Convention on the Rights of Persons with Disabilities, access to information and communication technologies including the web should be a basic human right [3]. In the United States, accessibility is required in design by national law according to Sect. 508 of the Rehabilitation Act of 1973. According to the World Wide Web Consortium (W3C), an international community that develops web standards, accessibility means that information can be obtained by all individuals regardless of hardware, software, language, location, or physical or mental capabilities [4]. In accordance with these standards, Web Content Accessibility Guidelines (WCAG 2.0) have been developed as part of the W3C to help ensure that web content is accessible.

There are a variety of web accessibility tools available that can assist designers in determining whether a website meets accessibility guidelines and standards [5]. Sun, Vu and Strybel [6] found that electronic content that was ranked high in accessibility resulted in better ratings of user experiences compared to electronic content ranked as low in accessibility. Sun, Manabat, et al. found that chapters from eBooks that were rated to be higher in accessibility were considered by participants to be more usable. Participants also indicated greater levels of satisfaction with books rated higher in accessibility compared to those rated lower in accessibility. Accuracy on reading comprehension questions was unaffected for participants without visual impairments. However, for visually impaired participants, accuracy was higher for books with non-STEM (science, technology, engineering or math) content compared to STEM content. Thus, the findings of Sun, Manabat, et al. demonstrated that web content that adheres to accessibility guidelines and standards could increase the subjective usability experience and performance of users with visual impairments.

According to the Nielson Norman Group, usability is an attribute that determines how easy an interface is to use, and is defined by learnability, efficiency, memorability, accuracy, and user satisfaction [7]. Usability is important regarding design in that if the interface or website is not usable, it will likely also be less accessible. Yesilada, Brajnik, Vigo, and Harper [8] found that accessibility standards were viewed as being highly related to usability. Although usability and accessibility are similar in design goals and approaches, the difference between the two is that usability does not focus specifically on individuals with restricted use due to disabilities or physical limitations. According to the Web Accessibility Initiative, the best method for developing websites is to incorporate both usability and accessibility standards and guidelines into the design process [4].

Websites that adhere to accessibility standards may also consider aspects of usability in the design process. Hallett, Dick, Jewett, and Vu [9] found that screen magnification techniques that utilized text wrapping (Responsive Web Design [RWD]) resulted in better user experience scores compared to standard screen magnification (SMS) software that did not allow for text wrapping. This is because SMS enlarges text onscreen in a manner that results in text becoming cutoff and requires horizontal scrolling which decreases usability [10]. Responsive web design wraps the enlarged text within the margins of the visible portion of the screen, making it more usable compared to SMS. Moreover, participants reported significantly higher usability ratings for the RWD than SMS condition, as measured by the System Usability Scale (SUS) [11]. Performance was measured by time-on-task and accuracy to reading comprehension questions. Although both methods (SMS and RWD) yielded similar accuracy scores on the comprehension questions, and received ratings that were within the acceptable range of the SUS, RWD had an average score equating to an "A," while SMS had an average score representative of a "C" [9]. Hallett et al.'s findings demonstrate that adhering to accessibility standards may improve the subjective usability of the system with which the interaction is occurring. Participants stated that they were more comfortable, and that it was more enjoyable to read the passages that used text wrapping (RWD condition). This is because RWD eliminated horizontal scrolling, which is a usability guideline recommendation.

Research has also demonstrated that the use of accessibility guidelines benefits not only individuals with disabilities but also individuals without disabilities. A recent study by Schmutz, Sonderegger, and Sauer [12] examined the level of adherence to accessibility standards and found that as more accessibility standards were implemented within a website, the percentage of tasks that were completed increases. Subjective measures obtained showed that participants gave higher ratings for usability, trustworthiness and aesthetics, and lower ratings of workload to more accessible sites [12]. These findings suggest that designing websites with high adherence to the WCAG 2.0 accessibility guidelines may benefit both users with and without disabilities.

With this in mind, it is important to study user groups that are not considered to be disabled but may have trouble in accessing content online. One such user group would be the elderly. According to the Web Accessibility Initiative, accessibility standards can also assist the elderly population who struggle with decreases in vision, hearing, and fine motor movements [13]. While it seems apparent that accessibility standards that deal with aspects of decreasing vision (color contrast, text size adjustment), fine motor movements (keyboard navigation), and hearing (closed captions, audio transcripts) can aid the elderly user in accessing web content, there has been little research demonstrating that high adherence to accessibility standards benefits the elderly. Aside from this, the implementation of accessibility standards into newly designed websites is slow moving, and in many cases, poorly integrated in the development process [4].

Yesilada et al. [8] found that those who are novices in the field of web accessibility may hold a popular view, not supported by experts and research, that implementing accessibility standards into the design of new websites is not worth the effort. Yesilada et al. found that the typical person feels that although web accessibility guidelines are beneficial to some users, they may not necessarily provide added benefit to users without disabilities, such as the elderly. Respondents were considered to be experts in the field of web accessibility if they spent at least 20% of their working hours dealing with accessibility, and have been working in the field of accessibility for seven or more years. Level of expertise was found to have an effect on responses where individuals with higher expertise agreed more with statements indicating that there was an added benefit to certain groups (such as those with low vision or the elderly) compared to responses from non-experts. These results show that, in general, the population with less expertise or experience in web accessibility may not see the necessity of following accessibility standards in web development.

Another view or idea that may be hindering the implementation of accessibility standards in designing websites is that by designing for accessibility, there could be a risk of creating usability issues for users without disabilities. However, research has shown this suggestion to be unfounded. Schmutz et al. [12] tested whether high adherence to accessibility standards would result in disadvantages to users without disabilities and found no observable differences. Yesilada, Brajnik, and Harper [14] had 76 college students evaluate websites according to accessibility standards to identify barriers that were common to both mobile device users and users with disabilities. They found that accessibility standards that help individuals with visual impairments and motor control issues access the web also help the standard user with accessibility of the web on a mobile device.

However, a study conducted by Hart, Chaparro, and Halcomb [15] found that design for accessibility by itself may not be beneficial. They conducted a study in two parts. The first part consisted of the identification of 40 websites that were checked against a set of heuristic design standards for accessibility. The second part consisted of participants browsing through three of those websites, ranging in adherence levels to accessibility standards, to discover any issues in the websites' usability. Hart et al. found that websites which had higher adherence to accessibility standards resulted in higher task success. However, they did not find any significant differences in task efficiency, user satisfaction, or preference from sites that were rated lower in adherence to accessibility standards. This suggests that although accessibility standards were implemented in the design of the websites, the usability of the website was overlooked. Curran, Walters, and Robinson [16] conducted an evaluation on various websites (including government, university, shopping, and charitable websites) to determine the level of adherence to accessibility standards and guidelines. Curran et al. found that although adherence to the accessibility guidelines is a matter that is becoming more important, many websites still fail to implement all recommendations. Moreover, most failed to even meet the most basic conformance level. This finding suggests that there is a lack of, or inadequate, implementation of accessibility guidelines, and demonstrates the importance of conducting usability and accessibility evaluations.

1.2 Age

The elderly, ranging in age from 50 and above, are the fastest growing group of internet users in society [15]. With age-related declines in vision and motor movement being experienced by this user group [17], it is important that accessibility not only aid the disabled, but also individuals such as the elderly who also have accessibility issues. Aside from the growing population of elderly internet users, there are still many elderly adults that do not use the internet. It is important to understand the reasons behind the disuse of the internet that still exists in the elderly population to determine if they can be overcome. Although the reasons for non-use vary, the top reasons for elderly individual's disuse of the internet (aside from cost of hardware and internet providers) include functional impairments and ergonomic barriers, such as small font sizes being hard to read due to visual deficits [17]. Websites that follow accessibility standards in design and that would allow for customization in such areas as text size adjustment could provide conditions that are beneficial to those who do not use the internet due to the lack of such capabilities.

Age-related issues involving vision are possibly the most common among the elderly. Fozard [18] found that declines in vision, visual acuity, and contrast sensitivity can begin to be noticed during the mid-40s. However, visual impairments typically occur at age 50, resulting in difficulty focusing on near objects, discriminating fine detail, or noticing changes in illumination. The ability of the older adults to differentiate between different hues of color, especially between blue-green, decreases because of reduced sensitivity to color [19–21]. Another visual factor that degrades as one ages is the ability to process visual information [17]. The effects of aging on the decline in cognitive functioning typically begin during midlife, but do not have a striking effect until age 70 and beyond, when very noticeable declines can occur [22, 23].

Another deficit that occurs as age increases is a decline in auditory functioning [17]. This can impede the use of electronic information by elderly internet users, especially when the information is presented in an audio format. The decline in auditory perception in the elderly is important to accessibility because information presented via sound, without the use of captioning or text transcripts, could be lost. This places a heavier reliance on visual information.

1.3 Screen/Font Size

It has been suggested that adults, 65 and older, may benefit from text that is displayed in less complex styles, such as sans serif, and with a font size in the 12- to 14-point range [24]. This suggestion was developed from the results of a series of surveys conducted across a 10-year period looking at the use of electronic devices by individuals 65 and older. Morrell and Echt [24] found that using a sans serif font type resulted in faster reading rates, and were reported as being easier to read by participants, as were larger character sizes between 12- to 14-point font. A study conducted by Ellis and Kurniawan [25] also found that the san serif font type, black Arial with a character size of 14, was preferred by elderly participants. Moreover, it was reported as being more appealing, and subjectively easier to read, compared to other font types such as Times New Roman. A more recent study conducted by Lin, Wu, and Cheng [26] found similar results to the study conducted by Ellis and Kurniawan [25]. However, screen size was also varied in Lin et al.'s study, using a color LCD e-reader. In this study, 60 Taiwanese students with normal or corrected-to-normal vision completed search tasks on pseudo-text. Their task was to identify target words as quickly and accurately as possible. Search time for words, accuracy, and visual fatigue were measured. Lin et al. found that search time for words varied significantly depending on screen size and character size. They found that as screen size increases from 6″ to 9.7″, when using 10- and 12-point font, search time significantly decreased. The same was observed for 8- and 14-point font size, but there was no significant change observed in search time from 8″ to 9.7″ screen size. In regards to the decrease in search time, Lin et al. found that 12-point font had the fastest word search time, followed closely by 14- and 10-, and finally 8-point font size. Lin et al. also found that screen size did not have a significant effect on accuracy, but font size did. They found that increasing the font size led to better accuracy, with 12- and 14-point font having the highest accuracy. It was also found that participants' reports of visual fatigue decreased as the size of the screen and font size increased. However, the results of this study were obtained using a younger population (i.e., high school students), and the results may not generalize to an older population.

Another study looking at the effect of screen size of mobile phones on college students' reading comprehension found that the screen size had no significant impact on user comprehension in terms of accuracy [27]. However, screen size did significantly impact reading time and ease of reading characters, in such that as screen size increased, the easier it became to read the characters. A study by Lee, Shieh, Jeng, and Shen [28] looked at the effect of font size on the legibility of electronic paper in search time for letters. College students completed a series of 12 conditions where they were required to scan alphanumeric pseudo text, and identify the target letter "A" as accurately and quickly as possible. Lee et al. found that as font size increased, the time it took to search for and locate desired information decreased. These findings suggest that the use of devices with smaller screens and font sizes may result in lower accuracy and longer search times for letters, unless the font size is able to be adjusted to a larger size.

1.4 Lighting/Contrast

The effects of age-related visual deficits regarding visual acuity are typically worsened in a dimly lit room. It has been found that even elderly individuals who had similar visual acuity to younger individuals in well-lit areas experienced significant reduction in visual acuity in low lit areas [29]. In a study that looked at the effect of lighting on the legibility of electronic paper, it was observed that accuracy increased as illumination increased [28]. A visual search task, conducted by Lee, Ko, Shen, and Chao [30], using letters as targets, looked at the effect of display type, light source, ambient illumination, interline spacing, and character size on visual performance and fatigue. Lee et al. found that search time decreases significantly, as illumination is increased from 300 to 1500 lx, for college participants who had completed visual scanning of three paragraphs for target letters. A study by Benedetto, Carbone, Drai-Zerbib, Pedrotti, and Baccino [31] looked at the effect of luminance and ambient illumination on visual fatigue and arousal during a reading task with college-aged participants. They found that elevated levels of light intensity from screen luminance reduce the number of eye blinks, which is one of the leading causes of visual fatigue. Benedetto et al. also found faster reading times for higher levels of luminance, but comprehension was not observed to be affected by the level of luminance or ambient illumination.

1.5 Visual Fatigue

As the use of computers and electronic devices have increased, so have a prevailing tendency of eye-related symptoms such as eyestrain, tired eyes, irritation, burning sensation, redness, blurred vision, and double vision [32]. These symptoms as relating to the use of computers and electronic devices have been termed Computer Vision Syndrome (CVS) [32]. The leading cause of CVS is dry eye brought on by the reduced tendency of the user to blink, which has also been shown to be related to elevated levels of illuminance [31, 32]. These symptoms result in the user experiencing visual fatigue. This is the state in which the visual mechanism, after operating in less than optimal conditions, ceases to function with maximum efficiency [33]. Visual fatigue does not primarily result from prolonged exposure or extensive involvement in reading, but from eye strain, which is greatly determined by illegibility due to inadequate illumination and color contrast [33].

According to Blehm et al. [32], there are several factors that could assist users in reducing the effects of visual fatigue. These factors include better lighting and less glare, optimal screen refresh rates, and proper screen positioning. In regards to lighting, reducing the intensity of lighting in the reading area will increase the readers blink rate, which in turn will reduce the level of dry eyes. The refresh rate of the screen on which individuals are reading from should be set at a high level to reduce the amount of flickering [32].

1.6 The Need for Customization

It is important to understand the difference between system driven personalization, user driven personalization, and customization. System driven personalization places control in the hands of the website. It delivers content and functionality to the user that matches his or her needs or interests based on users' past behaviors, without the user having to do anything [34]. An example of system driven personalization would be the ability of the website to make suggestions based on previous search inquires or interactions with the website. User driven personalization gives control to the user through customization, and allows the user to make changes to the physical layout, content, or functionality of the website to meet the individual's needs [35]. An example of this would be the ability of the user to change the color contrast of a website, adjust the size of the text displayed, or even move icons or images around on the website.

The idea of customization is important to website accessibility and usability. There are currently 56.7 million people in the United States who are classified as having a disability, which is roughly 1 out of every 5 people in our nation's population [36]. Allowing the user to have the ability to make customized changes to the web browser, or directly to the website, can greatly improve the accessibility of the web content to that individual [37]. In addition, adults can experience deficits in visual functioning as they age. There are over 28.8 million elderly internet users in the United States [36]. Ellis and Kurniawan [25] examined factors that should be taken into account in the development of a website specifically designed for use by the elderly population. They found that the most common features that participants wanted to customize were related to page visibility and legibility, or more specifically, font size, font type, and the color scheme. The ability to adjust text size, font type, and the color contrast can help older adults see well.

The goals of the present study were to determine if website customization for accessibility would lead to an increase in task performance and perceived usability of the system for older adults. In addition, the study examined whether customized text would reduce the reported level of visual fatigue when interacting with electronic content over a period of time. There has been little research on web customization, which can be due in part to the lack of platforms that allow the user to make customized changes to the webpage. A collaboration of departments at California State University Long Beach, along with the Knowbility organization, are developing a platform tool, called TRx. The TRx allows users to customize various aspects of the webpage to enhance the usability, accessibility, and user experience of the web. The TRx allows the user to make customized selections regarding: foreground-background color, font type and size, as well as, spacing parameters including line, letter, and word spacing. After the user selects the desired customization parameters, the TRx generates a stylesheet that can be incorporated into the internet browser, and it applies the desired selections.

This study examined the effects of customization and screen size on user performance, subjective usability, and reported levels of visual fatigue for online reading of older adults. We looked at performance in terms of reading time and accuracy, as measured by reading comprehension questions. Usability was determined by participants' responses to a SUS, and the level of user visual fatigue was assessed using a Visual Fatigue Questionnaire. We tested older adult participants who were between the ages of 50 to 69 years old. The ideal population would be users with visual disabilities, however, this population was selected due to the relative ease in obtaining participants for the study. Individuals with low vision have, in the past, been difficult to obtain for participation. The older adult participant sample for this study, although being one of convenience, was likely to have some visual problems associated with aging. We limited our older adults to participants less than 70 years old because of the cognitive decline typically found in participants 70 year or older [22, 23]. Since the reading passages are academic test passages, and not general reading materials, and the older adults may not be used to reading these types of passages, we ran a group of younger adult participants, who were between the ages of 18 to 22. The younger adults performed the reading tasks under customized and un-customized conditions to serve as baseline performance measures.

We developed a set of hypotheses based on previous research and they are: H1: Customization will lead to a decrease in reading time, but not in accuracy to reading comprehension, which has been shown to be insensitive [6, 9, 28]. H2: Customization will result in higher ratings of subjective usability. Allowing the user to customize webpage elements, such as, the ability to increase the font size, and select the font type, should result in higher user preferences. This has been shown in related studies conducted by Hallett et al. [9] and Ellis and Kurniawan [25], which were discussed previously. H3: Finally, customization is predicted to result in lower levels of reported visual fatigue from reading online content over a period of time. Previous research has supported this point in that increased font size, and better color contrast, has been shown to reduce reports of visual fatigue [26, 31, 33].

2 Methods

2.1 Participants

Thirty-four individuals participated in this study. Data for one participant in the older adult age group was excluded from the analysis and replaced due to extensive reading response times that were more than three standard deviations from the mean, for customized conditions, and more than two standard deviations from the mean for non-customized conditions. Data for one participant in the younger adult age group was excluded and replaced due to having low accuracy scores (20%) for the passage comprehension questions. Thus, the final data set consisted of data from 32 participants, 16 in each age group. All participants had normal or corrected-to-normal vision. All older adult participants indicated that they do not experience difficulty in typing or in using a mouse for cursor movement and selections. The older adults (5 males, 11 females) were 49 to 67 years old, with an average age of 57.5 years. The younger adults (1 male, 15 females) were 18 to 22 years old, with an average age of 18.6 years. Older adult participants were selected by means of convenience sampling, utilizing flyers and word-of-mouth recruitment. Younger adult participants were recruited by mean of the California State University Long Beach's Psychology 100 Subject Pool. Older adult participants received $15 for their participation in the study, while the younger adult participants received one credit towards their course research participation requirement.

Older adults were required to take the Mini-Mental State Examination to determine if any cognitive impairments were present that would interfere with the study's results. All results were in the normal range of cognitive functioning, with a mean score being greater than the 24-point criteria ($M = 28.94$, $SD = .99$). Older adult participants' computer proficiency was obtained from responses to the pretest demographics questionnaire. Responses were made based on a 5-point scale where participants rated their agreement to the statement, "I am proficient in using a computer," where 1 represented "Strongly Disagree" and 5 represented "Strongly agree". Older adults indicated that they were proficient in using the computer ($M = 4.06$, $SD = 1.39$). Older adult participants also indicated agreement with the statement that the use of a computer is essential in daily activities ($M = 4.12$, $SD = 1.36$), using a 1 to 5 scale. Both older and younger adult participants rated their level of proficiency in using the web browser Google Chrome on a scale of 1 to 7, with 1 representing "Not Proficient" and 7 representing "Very Proficient". Participants were all shown to be proficient in using Google Chrome (older adults: $M = 4.23$, $SD = 1.42$; younger adults: $M = 5.56$, $SD = 1.03$).

2.2 Materials

The current study was reviewed and approved by the university's Institutional Review Board (IRB). A standard Dell desktop computer, using a Windows 7 operating system, was used in this study. The monitor used was a Dell model U2312HM, with a 1920 × 1080 Pixel Resolution and 24-in. LCD display. The active screen size was scaled down to 7.5 in. for the small screen display conditions, and the full screen was used for the big screen display condition. The screen was not titled and was positioned perpendicular (90°) to the table. A Visual Fatigue Questionnaire was used to assess participants' levels of reported symptoms of fatigue. The questionnaire was administered before and after each reading condition completed during the study. Participants responded to the questionnaire by indicating their agreement to statements regarding visual fatigue symptoms on a scale that ranged from 1 (strong disagreement) to 5 (strong agreement). Although the scale was used in previous studies [9, 31], metrics regarding its reliability and validity were not available.

The TRx program was modified by the experimenters to allow participants to create customized selections of various elements of the reading material. Sample text was provided to each participant showing the effect of each customized selection made. Participants were given the opportunity, after viewing the sample text, to return and change any customized selection previously selected. The total time taken to complete all customized selections was less than ten minutes. A SUS was given to each participant to determine overall usability of the reading tasks with and without customization. The SUS consists of a 10-item questionnaire with a range of five response choices ranging from strongly disagree to strongly agree. Scores were normalized to create a percentile ranking. A score above 68 is considered above average, while anything below 68 out of 100 is below average [38]. A post experiment questionnaire was used to obtain participants' preferences relating to customization for both the big and small screen size, as well as responses relating to levels of visual fatigue.

A set of four reading passages, and two medical prescription passages, were used to assess participant performance in terms of accuracy to reading comprehension questions. The four reading passages average 557 words, and were obtained from online sixth-through eighth-grade standardized practice tests for reading comprehension. The medical prescription passages have 209 words for medicine A and 199 words for medicine B. The medical prescription passages have a seventh grade comprehension level, and were also obtained from online sixth- through eighth-grade standardized practice tests for reading comprehension. Participant reading times included the amount of time it takes to read each passage, plus the time it takes to answer the corresponding set of comprehension questions. Participants were allowed to refer to the reading passage when answering the comprehension questions, which removed any need for reliance on memory for passage content.

2.3 Design

The study employed a 2 (Customization: Customization vs. No Customization) × 2 (Screen Size: Big: 24 in. vs. Small: 7.5 in.) × 2 (Age Group: Younger vs. Older Adults) mixed design, with age group serving as the between-subjects variable, and customization and screen size being within-subjects variables. Half of the participants were randomly assigned to start in the customization condition first, and the other half started with the non-customization condition first. The order of the different passages was counterbalanced between participants. The dependent variables included performance (accuracy and time), subjective usability/preference ratings (SUS and Post Experiment Questionnaire) and reported levels of visual fatigue (Visual Fatigue Questionnaire).

2.4 Procedure

All participants were tested in the same lab environment. Participants were tested individually in a well-lit room with constant overhead lighting throughout the study. The lighting level did not produce any glare on the computer screen. When participants showed up for the study session, they were greeted and given an informed consent form. The informed consent included information regarding risks, benefits, and incentives associated with the study. After signing the consent form, participants were seated approximately 20 in. in front of the computer. The participant then completed a visual fatigue scale, which served as our baseline measure. Then they performed two reading tasks and one medical prescription passage within the customization or non-customization condition. Again, the condition the participant received first was counterbalance between participants. For each condition, the participants were given task instructions.

In the customized condition, the participants were given the opportunity to explore the TRx program. The experimenter showed the participants the different customization options and allowed them to select the setting that was most comfortable for them. Once all the customized settings were selected, the experimenter saved the settings and transferred the stylesheet containing the selections into Qualtrics. Qualtrics is an online survey platform that was used to display the reading passages and record the answers for the reading comprehension questions. Participants performed one reading passage within the customized condition with the big screen and one with the small screen. The order of screen size was counterbalanced between participants. After completing both reading passages, the participant was given the prescription task. Once the participant completed the prescription task, they were asked to fill out a paper-based VFS and SUS surveys. For the non-customized condition, the participants were given the task instructions and presented with the reading passages from Qualtrics. As with the customized condition, participants performed one reading task with the small screen and the other with the large screen (order counterbalanced between subjects) prior to performing the prescription task and filling out the paper based VFS and SUS.

After completing all experimental tasks, older adult participants were given the post experiment questionnaire. The post experiment questionnaire asked participants about their user preference for using customized text compared to non-customized text and level of agreement to statements relating to visual fatigue. The questionnaire consisted of a total of seven items: three to capture user preferences, three to capture users' agreement ratings with specific statements, and one open-ended question for user comments. Younger adults were not given the post questionnaire since we only intended to use their reading performance data as a baseline. Once the study was completed, participants were given a debriefing form and thanked for participation.

3 Results

All participants made customized selections in the customization condition. Each condition, customized and not customized, took approximately 20 min to complete with an overall task completion time of about 40 min.

3.1 Accuracy

Reading Task. A 2 (Customization vs. No Customization) × 2 (Screen size: Big vs. Small) × 2 (Age group: Younger Adults: vs. Older Adults) repeated measures ANOVA was performed. Age group was entered as the between-subjects variable, and accuracy on the reading passages was the dependent measure. Accuracy was determined by the percent of responses that a participant answered correctly (out of a total of five reading comprehension questions). Table 1 lists the mean accuracy for each condition by age group. There were no significant main effects found for customization, screen size, or age group. However, there was a significant interaction between age group and customization, $F(1, 30) = 8.62$, $MSE = 0.03$, $p < .01$, partial $\eta 2 = .22$ (Fig. 1). Tests of simple effects were performed by customization condition. For the customized condition, younger adults were more accurate (M = 91.9%, SD = 3.2%) compared to older adults (M = 76.9%, SD = 3.2%), $p = .002$. However, for the non-customized conditions, there was no significant difference in accuracy between older adults (M = 86.3%, SD = 3.4%) and younger adults (M = 82.5%, SD = 3.4%). Tests of simple effects by age group showed that younger adults tended to be more accurate in the customized condition compared to the non-customized condition, $p = .055$, but there was no difference in accuracy rates for older adults by condition, $p = .17$. No other significant interaction effects were found.

Table 1. Mean percent correct on reading comprehension passages by age group

Age group	Screen size	Customization	Reading accuracy	
			Mean	Standard error
Younger adults	Big	Customized	88.8%	5.1%
		Not-Customized	88.8%	4.8%
	Small	Customized	95.0%	3.9%
		Not-Customized	81.3%	4.9%
Older adults	Big	Customized	78.8%	5.1%
		Not-Customized	86.3%	4.8%
	Small	Customized	75.0%	3.9%
		Not-Customized	86.3%	4.9%

Prescription Task. A separate 2 (Customization vs. No Customization) × 2 (Age group: Young Adult: vs. Older Adults) repeated measures ANOVA was conducted, with age group serving as the between-subjects variable. The accuracy of responses to the prescription passages was the dependent measure. Overall, accuracy was high. The mean accuracy for young adults was 93.8% in both the customized and non-customized condition. The mean accuracy for older adults was 92.5% in the customized condition and 96.3% in the non-customized conditions. This ANOVA yielded no significant effects.

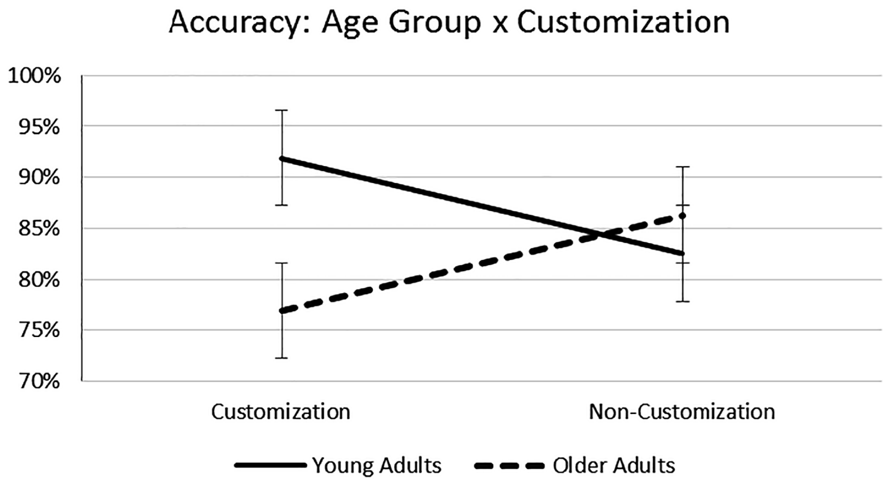

Fig. 1. Mean accuracy percentage. Accuracy in the customization condition was higher for younger adults compared to older adults whereas no differences were found for the non-customized condition.

3.2 Reading Time

Reading Task. A 2 (Customization vs. No Customization) × 2 (Screen size: Big vs. Small) × 2 (Age group: Young Adult: vs. Older Adults) repeated measures ANOVA was performed. Age group was the between-subjects variable, and reading time (reading passages and answering comprehension questions) was the dependent measure. The analysis did not yield any significant main effects or interaction effects (see Table 2 for means).

Table 2. Mean time on task (in Seconds) for reading comprehension passages by age group

Age group	Screen size	Customization	Reading time	
			Mean	Standard error
Younger adults	Big	Customized	267.87	14.67
		Not-Customized	282.31	15.03
	Small	Customized	268.25	17.24
		Not-Customized	270.94	16.91
Older adults	Big	Customized	284.00	14.67
		Not-Customized	262.06	15.03
	Small	Customized	296.81	17.24
		Not-Customized	278.75	16.91

Prescription Task. A separate 2 (Customization vs. No Customization) × 2 (Age group: Young Adult: vs. Older Adults) repeated measures ANOVA were performed.

Again, age group was the between-subjects variable, and reading response time for the two prescription passages was the dependent measure. There were no significant main effects or interactions. The average time for young adults was 144.94 s in the customized condition and 126.75 s in the non-customized condition, while for older adults it was 141.31 s in the customized condition and 118.25 s in the non-customized condition.

3.3 SUS

A 2 (Customized vs. Non-Customized) × 2 (Age group: Young Adult: vs. Older Adults) repeated measures ANOVA was performed to measure differences in subjective usability ratings. Age was the between-subjects variable, and the SUS score was the dependent measure. There were no significant main effects or interactions. All SUS scores were above the usability criteria for acceptable usability. The average SUS score for younger adults was 72.34 in the customized condition and 77.34 in the non-customized condition, while for older adults it was 76.56 in the customized condition and 76.87 in the non-customized condition.

3.4 Visual Fatigue

Visual fatigue was measured by an average score to Likert-like questions about visual fatigue, where 1 reflects low levels of visual fatigue, and 5 reflects high levels of visual fatigue. Since we were interested in how visual fatigue changed over the course of the study, a change score was obtained by subtracting both the customized and non-customized conditions visual fatigue averaged scores from the baseline score.

A 2 (Customized vs. Non-Customized) × 2 (Age group: Young Adult vs. Older Adults) repeated measures ANOVA was performed using the change in VFS score as the dependent measure. Age group served as the between-subjects variable. Overall, mean fatigue ratings were low, being less than 2, on a 5-point Likert Scale. There were no significant main effects or interactions.

3.5 Post Experiment Questionnaire

A post experiment questionnaire was given to participants in the older adult age group. For the first three items, participants indicated preference using a scale from 1 (using non-customized text) to 7 (using customized text), with 4 representing no preference. For the last three items, participants stated their level of agreement using a scale ranging from 1 (Strongly Disagree) to 7 (Strongly Agree), with 4 representing neither agreement nor disagreement. Individual one-sample T-tests were performed on Questions 1 to 3 to determine whether the preference for customization was significantly higher from the rating of 4, which represented no preference. Results indicated that participants' preference for using customization in the big screen conditions was significant, $t(15) = 3.67$, $p < 0.01$, while preference for using customization in the small screen, $t(15) = 1.78$, $p = 0.10$, and prescription, $t(15) = 1.85$, $p = 0.08$, conditions were non-significant (see Fig. 2). For Questions 4 to 6, descriptive analysis also showed a slight agreement that customized text decreased visual fatigue overall for both big and small screen reading tasks, as well as for the prescription tasks (see Table 3 for post experiment questionnaire means).

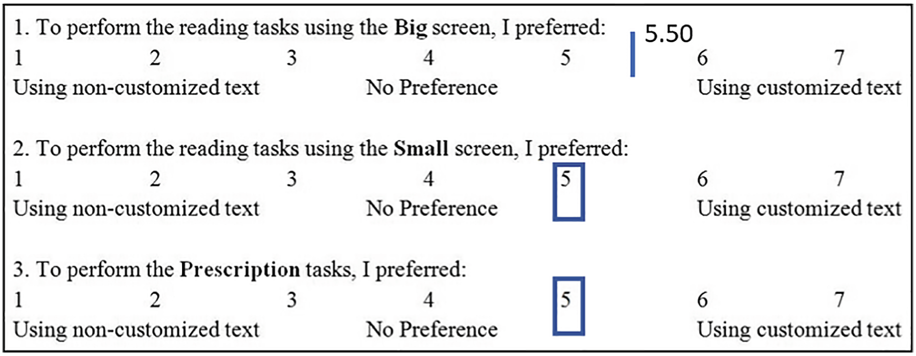

Fig. 2. Post experiment questionnaire customized preferences.

Table 3. Mean ratings given by older adults to the post experiment questionnaire

Older adult preference	Mean	Standard deviation
Big screen (preference for customization)	5.50	1.63
Small screen (preference for customization)	5.00	2.25
Prescription (preference for customization)	5.00	2.16
Decrease visual fatigue: big screen	5.38	1.63
Decrease visual fatigue: small screen	4.94	1.98
Decrease visual fatigue: prescription	4.88	1.75

4 Discussion

The present study evaluated whether reading tasks that utilized customized text elements would result in better performance, higher usability ratings, and decreased reports of visual fatigue for older adults. Performance was determined by accuracy (the percent of correct responses to reading comprehension questions) and reading/response time (the time it took participants to read each passage and answer the comprehension questions). The older adults' performance on the reading tasks were compared to those of younger adults.

The results showed only one significant interaction involving Customization and Age Group for accuracy on the reading passages. The customized reading passages resulted in higher accuracy rates for younger adults, but not older adults. In fact, older adult participants had numerically lower accuracy for the customized conditions compared to the younger adult participants. These findings do not support our hypothesis that customization will lead to a decrease in reading time but not in accuracy to reading comprehension questions. In fact, contrary to our hypothesis, we found that customization improved accuracy to reading comprehension questions for younger adults. Previous findings have shown that sans serif font types and font size between 12 and 14 point led to reduced reading times [24, 26]. When making the customized text selections participants were free to choose what they felt would be best for their online reading experience. However, after final customized selections were made, participants were not permitted to make changes to customized selection choices. Thus, customization selections were made only once and could not be optimized for the different screen sizes being tested.

For our younger adult group, 38% of participants customized their text to be 15 points and 43% of participants set their text to be 18 points or greater. For the older adult group, 43% of participants customized their text to be 15 points and 57% of participants set their text to be 18 points or greater. The lack of a difference in reading times may have resulted from the individual customized selections made by the participant. Larger text can be easier to read, but enlarging text can add to the reading task because participants will have to spend more time scrolling through the passages. Task demand was relatively low for the reading passages and comprehension questions. Pilot testing of the reading passages indicated that task completion times were all under 10 min for each passage and accuracy to the reading comprehension questions was above 80%. Because the sample size was small, and participants made different selections, there was not enough data available to meaningfully compare the reading times for each of the font sizes that were selected.

According to Tullis and Albert [11], SUS scores that are less than 50 are considered to be unacceptable, scores between 50 and 70 are considered to be marginally acceptable, and scores greater than 70 are considered to be acceptable in terms of usability. We found the SUS scores to be above 70 for all conditions. There were no significant differences in reported subjective usability observed between older adults and younger adults in either customized or non-customized reading conditions. Although our hypothesis stating that customization will result in higher rating of usability was not supported, both the customized and the non-customized reading conditions appeared to be within the acceptable range of usability according to participants from both age groups. The fact that customization did not reduce the usability ratings of the system is consistent with previous findings that showed designing for accessibility does not reduce system usability [12, 14].

Although previous studies have shown that visual fatigue may be decreased by increasing font size and improving color contrasts [26, 32, 33], there were no significant differences in participants' reported levels of visual fatigue found in the current study. These results do not support our hypothesis that customization will result in lower levels of reported visual fatigue. The changes in visual fatigue observed in the present study were small. The visual fatigue questionnaire used in the current study consisted of questions with a 5-point rating scale, which had a minimum score of 1 (representing low visual fatigue) and a maximum score of 5 (high visual fatigue). Baseline visual fatigue levels suggested that participants were not fatigued at the beginning of the study (mean visual fatigue scores were below a score of 2). Participants did not report significant increases in visual fatigue levels by the end of the study, and the means for all conditions were still under 2.0. Chi and Lin [39] found that changes in visual fatigue, as measured by the visual fatigue scale, were evident for reading tasks lasting about 60 min. The participants in the current study were able to finish the study within an hour. Thus, the time spent on the reading task may not have been long enough for detection of visual fatigue. In fact, in the post questionnaire, older adults only showed moderate levels of agreement with statements indicating that the customized conditions decreased their visual fatigue.

The lack of significant findings of the current study may be a result of the small sample size, which was limited by the availability of older adults. The observed power for all non-significant effects was between the range of 0.05 to 0.375, which reflects low power. Further research with a larger sample size is needed to increase the power, and to determine if the current results are representative.

Prior research has shown that to improve the accessibility and usability of the online experience, websites should be designed with the user in mind and allow for customizations of text elements including font family and size, color contrast, and spacing parameters [25]. Websites that incorporate the ability to customize are more likely to aid individuals with disabilities or populations without disabilities, such as older adults, that may experience difficulty in accessing electronic information. Based on findings from the current study, designing websites that allow for customization may lead to an increase in the younger population's ability to accurately answer reading comprehension questions. However, allowing for unlimited combinations of color for foreground and background text might not be an optimal design for electronic text. Further research should be conducted to determine what color combinations would work best for customized electronic text. Future research should also look at predetermined selections regarding font size, screen size, and color combinations to determine if allowing users to make unrestricted customized selections is more beneficial compared to the use of preset selections.

Another limitation of the current study is that, as noted earlier, participants were not able to make changes to the customized selections after they started the reading task in the customized condition. In practice, users may set an initial customize scheme, and then change the settings as they perform the task. That is, users may like a certain font size or color contrast setting at the start of the task, but after performing the task, they may want to make adjustments. Users may have also wanted to select different customization schemes based on the small versus large screen. Although the current implementation of the TRx customization software allowed users to see the changes in the text as a function of the customized selections, once the customized settings were finalized, the experimenter had to generate code for the selections and upload the code into Qualtrics. This process did not permit adjustments to be made for a single aspect in real time. It could be the case that a platform that allows users to customize in real time would lead to better performance, and future research should explore this possibility.

Acknowledgments. We thank the students in the Center for Usability in Design and Accessibility (CUDA) for help in proof reading and formatting the paper.

References

1. Mayes, D.K., Sims, V.K., Koonce, J.M.: Comprehension and workload differences for VDT and paper-based reading. Int. J. Industr. Ergon. **28**, 367–378 (2001)
2. Porion, A., Aparicio, X., Megalakaki, O., Robert, A., Baccino, T.: The impact of paper-based versus computerized presentation on text comprehension and memorization. Comput. Hum. Behav. **54**, 569–576 (2016)
3. United Nations. Convention on the Rights of Persons with Disabilities: Articles (2007). https://www.un.org/development/desa/disabilities/convention-on-the-rights-of-persons-with-disabilities/convention-on-the-rights-of-persons-with-disabilities-2.html
4. World Wide Web Consortium (W3C). Accessibility (2016). https://www.w3.org/standards/webdesign/accessibility
5. Sun, Y.T., Manabat, A.K., Chan, M.L., Chong, I., Vu, K.-P.L.: Accessibility evaluation: manual development and tool selection for evaluating accessibility of e-textbooks. In: Hale, K., Stanney, K. (eds.) Advances in Neuroergonomics and Cognitive Engineering. AISC, vol. 488, pp. 327–337. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-41691-5_28
6. Sun, Y.T., Vu, K.-P.L., Strybel, T.Z.: A validation test of an accessibility evaluation method. In: Ahram, T., Falcão, C. (eds.) AHFE 2017. AISC, vol. 607, pp. 625–633. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-60492-3_59
7. Nielson, J.: Usability 101: Introduction to usability (2012). https://www.nngroup.com/articles/usability-101-introduction-to-usability/
8. Yesilada, Y., Brajnik, G., Vigo, M., Harper, S.: Exploring perceptions of web accessibility: a survey approach. Behav. Inf. Technol. **34**(2), 119–134 (2013)
9. Hallett, E.C., Dick, W., Jewett, T., Vu, K.-P.L.: How screen magnification with and without word-wrapping affects the user experience of adults with low vision. In: Ahram, T., Falcão, C. (eds.) AHFE 2017. AISC, vol. 607, pp. 665–674. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-60492-3_63
10. Sherwin, K.: Beware horizontal scrolling and mimicking swipe on desktop. Nielsen Norman Group (2014). https://www.nngroup.com/articles/horizontal-scrolling/
11. Tullis, T., Albert, B.: Measuring the User Experience: Collecting, Analyzing, and Presenting Usability Metric, 2nd edn. Morgan Kaufmann, Waltham (2013)
12. Schmutz, S., Sonderegger, A., Sauer, J.: Implementing recommendations from web accessibility guidelines: would they also provide benefits to nondisabled users. Hum. Factors **58**(4), 611–629 (2016)
13. World Wide Web Consortium (W3C). Social factors in developing a web accessibility business case for your organization (2012). https://www.w3.org/WAI/bcase/soc#groups
14. Yesilada, Y., Brajnik, G., Harper, S.: Barriers common to mobile and disabled web users. Interact. Comput. **23**(5), 525–542 (2011)
15. Hart, T.A., Chaparro, B.S., Halcomb, C.G.: Evaluating websites for older adults: adherence to 'senior-friendly' guidelines and end-user performance. Behav. Inf. Technol. **27**(3), 191–199 (2008)
16. Curran, K., Walters, N., Robinson, D.: Investigating the problems faced by older adults and people with disabilities in online environments. Behav. Inf. Technol. **26**(6), 447–453 (2007)
17. Hawthorn, D.: Possible implications of aging for interface designers. Interact. Comput. **12**(5), 507–528 (2000)
18. Fozard, J.: Vision and hearing in aging. In: Birren, J., Sloane, R., Cohen, G. (eds.) Handbook of Mental Health and Aging, pp. 150–170. Academic Press, San Diego (1990)

19. Echt, K.: Designing web based health information for older adults: visual considerations and design directives. In: Older Adults. Health Information and the World Wide Web, pp. 61–88 (2002)
20. Helve, J., Krause, U.: The influence of age on performance in the Panel-D15 colour vision test. Acta Opthalmologica **50**, 896–901 (1972)
21. Weale, R.: Retinal illumination and age. Trans. Illum. Eng. Soc. **26**, 95–100 (1961)
22. Aartsen, M.J., Smits, C.H., van Tilburg, T., Knipscheer, K.C., Deeg, D.J.: Activity in older adults: cause or consequence of cognitive functioning? A longitudinal study on everyday activities and cognitive performance in older adults. J. Gerontol. **57**(2), 153–162 (2002)
23. Vestergaard, S., Thinggaard, M., Jeune, B., Vaupel, J.W., McGue, M., Christensen, K.: Physical and mental decline and yet rather happy? A study of Danes aged 45 and older. Aging Mental Health **19**(5), 400–408 (2015)
24. Morrell, Q., Echt, K.: Designing written instructions for older adults learning to use computers. In: Fisk, A., Rogers, W. (eds.) Handbook of Human Factors and the Older Adult, pp. 335–361. Academic Press, San Diego (1996)
25. Ellis, D.R., Kurniawan, S.H.: Increasing the usability of online information for older users: a case study in participatory design. Int. J. Hum.-Comput. Interact. **12**(2), 263–276 (2000)
26. Lin, H., Wu, F.-G., Cheng, Y.-Y.: Legibility and visual fatigue affected by text direction, screen size and character size on color LCD e-reader. Displays **34**, 49–58 (2013)
27. Ghamdi, E.A., et al.: The effect of screen size on mobile phone user comprehension of health information and application structure: an experimental approach. J. Med. Syst. **40**(1), 11:1–11:18 (2016)
28. Lee, D.-S., Shieh, K.-K., Jeng, S.-C., Shen, I.-H.: Effect of character size and lighting on legibility of electronic papers. Displays **29**, 10–17 (2008)
29. Kline, D., Scialfa, C.: Sensory and perceptual functioning: basic research and human factors implications. In: Fisk, A., Rogers, W. (eds.) Handbook of Human Factors and the Older Adult, pp. 27–54. Academic Press, San Diego (1996)
30. Lee, D.-S., Ko, Y.-H., Shen, I.-H., Chao, C.-Y.: Effect of light source, ambient illumination, character size and interline spacing on visual performance and visual fatigue with electronic paper displays. Displays **32**, 1–7 (2011)
31. Benedetto, S., Carbone, A., Drai-Zerbib, V., Pedrotti, M., Baccino, T.: Effects of luminance and illuminance on visual fatigue and arousal during digital reading. Comput. Hum. Behav. **41**, 112–119 (2014)
32. Blehm, C., Vishnu, S., Khattak, A., Mitra, S., Yee, R.W.: Computer vision syndrome: a review. Surv. Ophthalmol. **50**(3), 253–262 (2005)
33. Demilia, L.A.: Visual fatigue and reading. J. Educ. **151**(2), 4–34 (1968)
34. Schade, A.: Customization vs. personalization in the user experience. Nielsen Norman Group (2016). https://www.nngroup.com/articles/customization-personalization/
35. World Wide Web Consortium (W3C). Personalization Semantics Explainer 1.0 (2018). https://www.w3.org/TR/personalization-semantics-1.0/
36. U.S. Census Bureau. Fact for features: older Americans, May 2016. https://www.census.gov/newsroom/facts-for-features/2016/cb16-ff08.html
37. World Wide Web Consortium (W3C). Better web browsing: tips for customizing your computer. Web Accessibility Initiative (2010). https://www.w3.org/WAI/users/browsing
38. U.S. Department of Health and Human Services. System Usability Scale. Usability.gov (2017). https://www.usability.gov/how-to-and-tools/methods/system-usability-scale.html
39. Chi, C., Lin, F.: A comparison of seven visual fatigue assessment techniques in three data-acquisition VDT tasks. Hum. Factors **40**(4), 577–590 (1998)