

Implementing Recommendations From Web Accessibility Guidelines: Would They Also Provide Benefits to Nondisabled Users

Sven Schmutz, Andreas Sonderegger, and Juergen Sauer, University of Fribourg, Switzerland

Objective: We examined the consequences of implementing Web accessibility guidelines for nondisabled users.

Background: Although there are Web accessibility guidelines for people with disabilities available, they are rarely used in practice, partly due to the fact that practitioners believe that such guidelines provide no benefits, or even have negative consequences, for nondisabled people, who represent the main user group of Web sites. Despite these concerns, there is a lack of empirical research on the effects of current Web accessibility guidelines on nondisabled users.

Method: Sixty-one nondisabled participants used one of three Web sites differing in levels of accessibility (high, low, and very low). Accessibility levels were determined by following established Web accessibility guidelines (WCAG 2.0). A broad methodological approach was used, including performance measures (e.g., task completion time) and user ratings (e.g., perceived usability).

Results: A high level of Web accessibility led to better performance (i.e., task completion time and task completion rate) than low or very low accessibility. Likewise, high Web accessibility improved user ratings (i.e., perceived usability, aesthetics, workload, and trustworthiness) compared to low or very low Web accessibility. There was no difference between the very low and low Web accessibility conditions for any of the outcome measures.

Conclusion: Contrary to some concerns in the literature and among practitioners, high conformance with Web accessibility guidelines may provide benefits to users without disabilities.

Application: The findings may encourage more practitioners to implement WCAG 2.0 for the benefit of users with disabilities and nondisabled users.

Keywords: Web accessibility, nondisabled users, WCAG 2.0, performance, usability

INTRODUCTION

People with disabilities may face various barriers in their daily activities. For example, a wheelchair user is not able to move up a flight of stairs or a person who has no speech cannot answer a phone. An important activity that may also entail barriers for people with disabilities is the use of the World Wide Web (Web). For instance, hand tremor can make it difficult to click on a link of small size, audio content may not be accessible due to deafness, or low text-to-background contrast cannot be read by users with visual impairments (Vu & Proctor, 2011). About 15% of the world's population has some kind of disability (World Health Organization, 2011), and many of these disabilities result in difficulties in using Web sites (e.g., cognitive, hearing, motor, and visual impairments; Ruth-Janneck, 2011). Thus, a substantial portion of people has restricted or no access to information on Web sites, resulting in considerable disadvantages for the people concerned because of the Web's pervasiveness and importance in society.

To overcome this issue, Web accessibility (hereafter "accessibility") aims to ensure that "people with disabilities can perceive, understand, navigate, and interact with the web" (Henry, 2006, p. 2). A typical measure for achieving this objective is the use of guidelines for accessible Web design (accessibility guidelines), which recommend specific Web site characteristics to support users with disabilities. For instance, they recommend using text alternatives for audio content to support users with hearing impairments or suggest a minimum contrast between text and background to support users with visual impairments in reading text (Caldwell, Cooper, Reid, & Vanderheiden, 2008). Although authors of some studies examined the validity of accessibility guidelines in terms of effects on users with disabilities (e.g., Power, Freire, Petrie, & Swallow, 2012; Rømen &

Address correspondence to Sven Schmutz, Department of Psychology, University of Fribourg, Rue de Faucigny 2, 1700 Fribourg, Switzerland; e-mail: sven.schmutz@unifr.ch.

HUMAN FACTORS

Vol. 58, No. 4, June 2016, pp. 611–629

DOI: 10.1177/0018720816640962

Copyright © 2016, Human Factors and Ergonomics Society.

Svanaes, 2012; Ruth-Janneck, 2011), little research has focused on effects of implementing accessibility guidelines for nondisabled users (Yesilada, Brajnik, Vigo, & Harper, 2013). Improving our knowledge of their effects on nondisabled users is important because nondisabled users also use accessible-designed Web sites and even represent the vast majority of users. A crucial goal in Web site design is to satisfy as many users as possible (Vu & Proctor, 2011). Consequently, adverse effects on nondisabled users may hinder the implementation of accessibility guidelines in practice. Conversely, positive side effects on nondisabled users may encourage practitioners to use accessibility guidelines. Therefore, we aim to present empirical evidence of the consequences of implementing accessibility guidelines for nondisabled users.

Web Content Accessibility Guidelines

Web content accessibility guidelines are a tool that can be used for creating content considering the needs of people with disabilities or for evaluating a Web site concerning accessibility (Chisholm & Henry, 2005). The most commonly used set of accessibility standards is the Web Content Accessibility Guidelines 2.0 (WCAG 2.0; Caldwell et al., 2008). It has found its way into the laws of several countries, such as Australia, Canada, Germany, Japan, and Hong Kong (Rogers, 2015), and constitutes the International Organization for Standardization's (2012) standard for accessibility. The WCAG 2.0 comprises 61 success criteria, which provide specific recommendations on how to make Web content accessible to users with different impairments (e.g., cognitive, hearing, visual, and physical impairments). For instance, all functionalities should be available from the keyboard, or captions should be used to describe audio content. Each criterion can be tested for whether it is satisfied. According to the degree of conformance with these criteria, Web sites can be classified into one of three categories: low accessibility (A), high accessibility (AA), and highest accessibility (AAA) (see Caldwell et al., 2008, for details). According to WCAG 2.0, Level A or higher may be considered accessible. Hence, we will term Web sites nonaccessible (NA) if they offer lower conformance than Level A.

Rationale for Implementing Accessibility

The application of WCAG 2.0 in practice is still rare. In recent studies, more than 95% of Web sites investigated were classified as NA (e.g., Gonçalves, Martins, Pereira, Oliveira, & Ferreira, 2013; Nurmela, Pirhonen, & Salminen, 2013). Among practitioners, reasons for implementing accessibility have been market benefits, legal requirements, the intention to be inclusive, and to design better products (Farrelly, 2011; Loiacono & Djasmasbi, 2013; Yesilada, Brajnik, Vigo, & Harper, 2012). Practitioners also reported reasons hindering accessibility implementation, including lack of financial benefits and no demand of clients and management (Farrelly, 2011; Freire, Russo, & Fortes 2008; Lazar, Dudley-Sponaugle, & Greenidge, 2004). Another issue that may prevent practitioners from applying accessibility is prevailing negative beliefs about accessible design. Such beliefs have already been extensively discussed in the literature (see Ellcessor, 2014, for a review). This includes the beliefs that accessible Web sites are ugly and boring (e.g., Lawson, 2006; Petrie, Hamilton, & King, 2004) and that accessibility provides benefits to only a small number of users (e.g., Mlynarczyk, 2012).

Against this background, knowing more about effects of accessibility on nondisabled users may be important for decisions on implementing accessibility in practice. Particularly, positive effects of accessibility on nondisabled users would be in line with reasons for implementing accessibility (e.g., market benefits or being inclusive) because a wide range of users would be positively affected. In addition, since practitioners mention legal requirements to be an important reason for implementing accessibility (Loiacono & Djasmasbi, 2013), it is to be expected that the number of Web sites conforming to WCAG 2.0 will grow due to the dissemination of these guidelines by means of different national laws (e.g., Murphy, 2013). As this growth will lead to an increase in the number of accessible Web sites available to nondisabled users, it is important to investigate possible side effects of accessible design on this user group. Given these considerations, previous studies suggest a need to empirically examine

the consequences of implementing accessibility in Web sites for nondisabled users (e.g., Yesilada et al., 2012, 2013).

Research on Accessibility With Nondisabled Users

One of the first studies that stated a relation between accessibility and nondisabled users focused on visual design of Web sites (Petrie et al., 2004). Fifty-one users with disabilities evaluated the accessibility of 100 Web sites by reporting problems encountered. The authors discussed the problems identified in regard to visual design, concluding that accessibility referred to aspects of visual design that may also pertain to nondisabled users (e.g., color contrast or visual structure). In line with this conclusion, Mbipom (2009) showed that Web sites with high ratings on certain dimensions of aesthetics (i.e., being clear, clean, and organized) violated fewer WCAG 1.0 criteria (first version of the WCAG; <http://www.w3.org/TR/WCAG10/>) than Web sites with lower ratings of aesthetics. Another study compared users who were blind to nondisabled users regarding task completion rate and time (Disability Rights Commission, 2004). Both user groups solved tasks on three Web sites with high accessibility ratings (according to WCAG 1.0) and three Web sites with low ratings. Whereas users who were blind solved more tasks on highly accessible Web sites, nondisabled users' task completion rate was not affected by accessibility level. Of particular interest was the result that both user groups were faster using highly accessible Web sites compared to Web sites with low accessibility.

Further work suggested that users with disabilities and nondisabled users may encounter similar problems but that the impact may be stronger on users with disabilities (Petrie & Kheir, 2007). In Petrie and Kheir's (2007) study, six users who were blind and six nondisabled users were compared regarding problems in using Web sites. All participants solved tasks on two Web sites and were asked to report occurring problems. Results showed that about 15% of the problems reported were encountered by both user groups. Although the overlap appears to be rather small, it confirmed that users with and without disabilities may be

affected by the same Web site characteristics. Further research showed that high WCAG 1.0 conformance led to higher usability ratings by nondisabled users compared to low conformance (Huber & Vitouch, 2008) and also to higher scores in automated usability testing (Sullivan & Matson, 2000).

However, there is also work that did not show a relationship between WCAG 1.0 conformance and subjective ratings or performance of nondisabled users (Arrue, Fajardo, Lopez, & Vigo, 2007). Although the vast majority of studies were based on WCAG 1.0, there is also recent work that made use of the improved guidelines, WCAG 2.0, when examining nondisabled users (Pascual, Ribera, Granollers, & Coiduras, 2014). They compared an A Web site to an NA Web site with regard to performance and subjective measures. The sample consisted of four nondisabled users and nine visually impaired users. For nondisabled users, all performance measures and subjective measures showed similar means for the A Web site and the NA Web site. Users with visual impairments completed more tasks on the A Web site than on the NA Web site, and they were also faster in doing so.

Similar to the approach of the present work, authors of research in a related field (i.e., designing for older adults) investigated possible positive side effects of design recommendations. It emerged that designing for older adults may also benefit younger adults (e.g., Chadwick-Dias, McNulty, & Tullis, 2003; Johnson & Kent, 2007; Pak & Price, 2008; Westerman, Davies, Glendon, Stammers, & Matthews, 1995). Furthermore, Web site design recommendations for older adults (e.g., Badre, 2002; Mead, Lamson, & Rogers, 2002; Rogers & Fisk, 2001) overlap considerably with the recommendations of WCAG 2.0 (e.g., high contrast, intuitive link texts, consistent design, left-aligned text). Given the overlap between the design for older adults and accessibility recommendations, and considering the fact that design for older adults will also provide benefits to younger adults, it is conceivable that accessibility will also have positive side effects on nondisabled users.

Overall, previous research reflected a positive influence of accessibility for nondisabled users. The studies reviewed provide important first insights into the effects of accessibility on

nondisabled users, but there are still some knowledge gaps. First, note that all studies reported (except Pascual et al., 2014) were conducted before the release of WCAG 2.0 in 2008, being based on WCAG 1.0 as a reference standard. This time frame is important because WCAG 2.0 differs considerably from WCAG 1.0 (Reid & Snow-Weaver, 2008, 2009). Hence, research results based on WCAG 1.0 need to be treated with some caution. Second, most studies included rather small samples of nondisabled users (e.g., Pascual et al., 2014, $n = 4$; Petrie & Kheir, 2007, $n = 6$) or did not include nondisabled users (e.g., Petrie et al., 2004; Sullivan & Matson, 2000). This composition suggests a strong need for studies with larger samples of nondisabled users. Third, studies on performance (e.g., task completion time) of nondisabled users mainly reported descriptive statistics (i.e., means or percentages) and conducted no inferential statistical tests (e.g., Disability Rights Commission, 2004; Pascual et al., 2014). Such inferential statistical analyses should be conducted. Fourth, most studies emphasized ecological validity by comparing various real Web sites with different levels of accessibility (e.g., Pascual et al., 2014; Petrie & Kheir, 2007). Although studies with high ecological validity are important for practical application, researchers also need to use more controlled experiments to gain further insights into the effects of accessibility on nondisabled users. Fifth, until now, none of the studies that addressed accessibility and nondisabled users had measured a broader range of outcome variables, using both objective measures (e.g., task completion time) and subjective ones (e.g., perceived workload). Such a broader measurement approach may help gain a more comprehensive understanding of the influence of accessibility on nondisabled users.

The Present Study

The goal of this experiment was to examine the consequences of implementing recommendations from current accessibility guidelines (i.e., WCAG 2.0) for nondisabled users. As an independent variable, accessibility was manipulated by modifying 13 WCAG 2.0 recommendations in an existing municipal Web site, resulting in three versions of the Web site with

different levels of accessibility: Levels AA, A, and NA. Level AAA was not included because of its rare prevalence in practice (e.g., Nurmela et al., 2013). Apart from WCAG 2.0 conformance, the Web sites were identical (e.g., same content and same number of menu items). The Web sites were evaluated by means of a usability test, taking a range of performance measures (task completion time and task completion rate) and subjective ratings (usability, aesthetics, trustworthiness, affect, and workload).

Based on the literature review, beneficial effects of accessibility on performance and subjective measures were expected for nondisabled users. More specifically, it was predicted that Level AA and Level A would lead to lower task completion time, higher task completion rate, and higher perceived usability than Level NA. Furthermore, it was assumed that Level AA and Level A would result in higher ratings of aesthetics, trustworthiness, and positive affect but lower ratings of negative affect and workload than Level NA. Overall, the more the Web site corresponded to the guidelines, the higher we expected ratings and performance to be (i.e., $NA < A < AA$).

METHOD

Participants and Design

Sixty-one participants took part in the study (see Table 1 for details). Participants were students or recent graduates. They were unpaid but students received credits for their participation. We required participants to have normal acuity (with or without glasses) and color vision (i.e., no visual impairment diagnosed by physician or registered by health insurance). We relied on self-report data from participants, and no vision test was conducted.

The study employed a one-factorial between-subjects design, in which a Web site was manipulated at three levels of WCAG 2.0 conformance: Level NA (very low conformance), Level A (low conformance), and Level AA (high conformance).

The Web Sites

Content and characteristics. The Web sites are based on a municipal Web site of a town in

TABLE 1: Overview of the Sample for the Three Accessibility Conditions

Sample characteristic	Total	Accessibility Level (WCAG 2.0)			ANOVA	
		NA Web Site	A Web Site	AA Web Site	F(df)	p
Participants: <i>n</i>	61	21	20	20		
Women: <i>n</i> (%)	48 (78.7)	13 (61.9)	17 (85.0)	18 (90.0)		
Age: <i>M</i> (<i>SD</i>)	21.4 (2.4)	21.8 (3.6)	21.5 (1.5)	21.1 (1.6)	0.5 (2, 60)	.62
Education in years: <i>M</i> (<i>SD</i>)	12.5 (1.4)	12.6 (1.6)	12.0 (0.0)	12.8 (1.8)	1.6 (2, 60)	.21

Note. WCAG = Web Content Accessibility Guidelines; NA = very low conformance; A = low conformance; AA = high conformance.

the country of Liechtenstein (www.eschen.li). This Web site is certified as AA according to WCAG 2.0. The Web site contains mainly text information, pictures, and a contact form. There is no multimedia content available, such as video (the Web cam and map function provided by the original Web site were removed for testing). Furthermore, the Web site is primarily based on html as well as CSS and contains little JavaScript. The size of the original Web site was slightly reduced (i.e., seven pages of no relevance for the present study were removed), and three copies of the Web site were made. The design of one Web site remained similar to the original corresponding to WCAG 2.0 Level AA. The two further copies were adapted according to Level A and Level NA (see Figure 1 for screenshots).

Web Site Manipulation

Several steps were taken to obtain manipulations corresponding to Level A and Level NA. First we had initial discussions with three accessibility experts about success criteria relevant to practitioners and possible manipulations of the Web site. These experts were Web developers with several years of work experience in accessibility and design and members of the nonprofit organization Access for All. Access for All is the competence center for accessibility in Switzerland and offers a certification of Web sites according to WCAG 2.0. Second, we aimed to identify frequent violations of WCAG 2.0 recommendations by reviewing the literature and the document *Understanding WCAG 2.0* (Cooper, Kirkpatrick, & O Connor, 2014) and by

screening 500 municipal Web sites. This process resulted in 10 Web sites’ characteristics being chosen that were considered relevant for the present study. For the A Web site, the following characteristics were manipulated: contrast, text alignment, precision of link description, appropriateness of headings, focus visibility, number of section headings, and consistency in link style. Additionally, the following characteristics were manipulated for the NA Web site: precision of form description, focus order, and error identification (see the appendix for details). Third, we had repeated discussions with the experts on how to adapt and implement common failures in Web sites for the low and very low accessibility conditions. Note that the chosen criteria were considered to be typical accessibility characteristics of particular relevance based on earlier empirical work (e.g., Freire, Petrie, & Power, 2011; Power et al., 2012; Rello, Kanvinde, & Baeza-Yates, 2012; Ruth-Janneck, 2011). A further reason for choosing these criteria was that most of the criteria were of general relevance because it has been shown that they also provide benefits to other user groups, such as older users (e.g., Chadwick-Dias et al., 2003; Johnson & Kent, 2007; Kurniawan & Zaphiris, 2005; Nayak, Priest, Stuart-Hamilton, & White, 2006; Sayago, Camacho, & Blat, 2009).

Manipulation Check

In order to validate the manipulations of the Web sites, we did additional testing by using a synchronous remote method (using screen-sharing technology called TeamViewer; cf. Andreasen, Nielsen, Schröder, & Stage, 2007; Dray & Siegel,

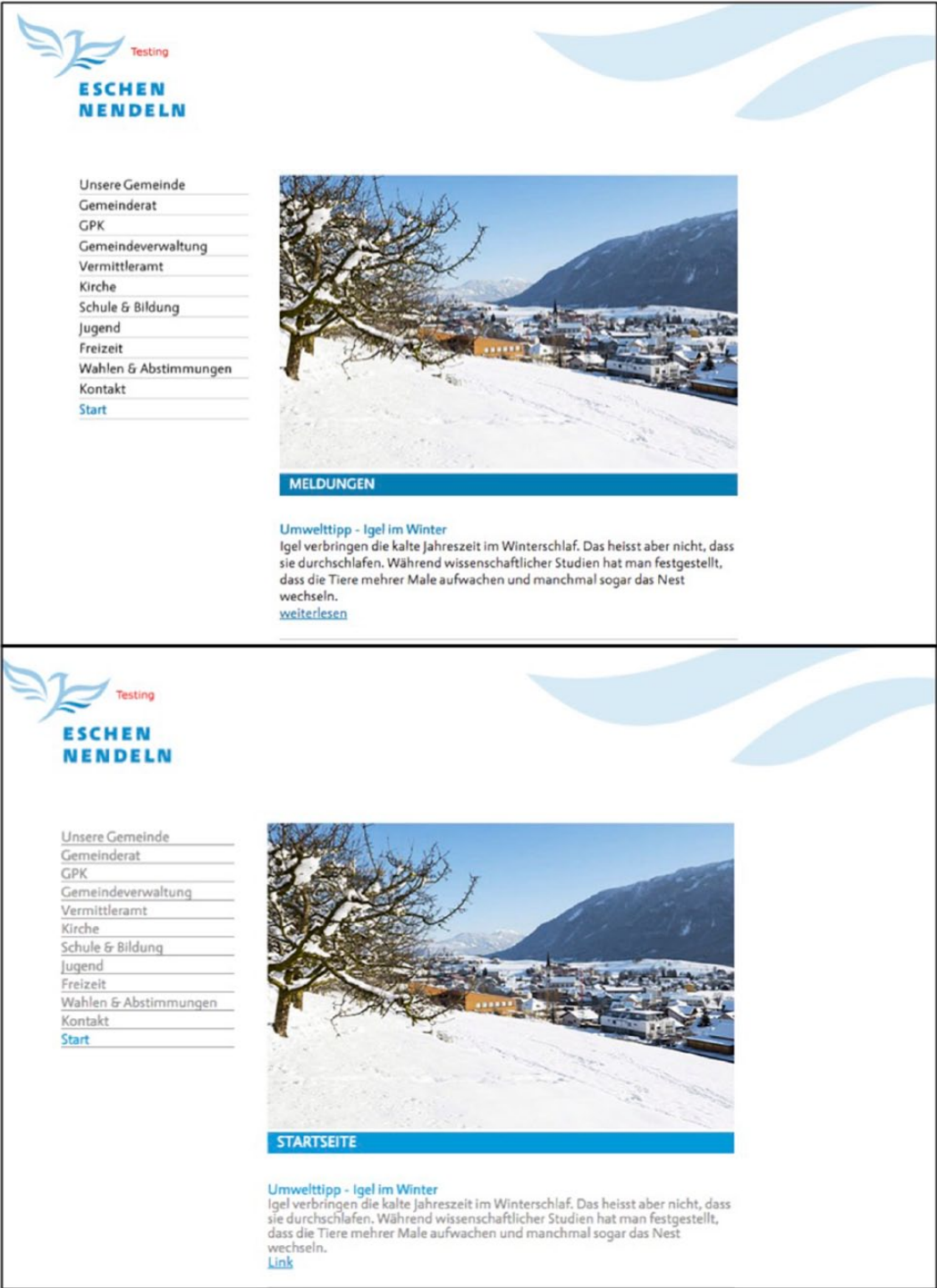


Figure 1. Screenshots of the home page of the Web sites used for testing (top, Level AA [high conformance]; bottom, Level NA [very low conformance]).

TABLE 2: Tasks to Be Completed on the Web Site

Task	Description
1	Identify the bus service that will take you to the sports center.
2	Fill out a form to order firewood from the municipality.
3	Find the e-mail address of the person responsible for education in town.
4	Identify the motorway that connects this town with Switzerland.
5	Find a phone number to book the barbecue area for a group event.

2004). In a two-factorial between-subjects design, 55 users without visual impairments (age, $M = 45.1$, $SD = 14.8$, range = 22–71) and 55 users with visual impairments (i.e., a maximum eyesight of 20% on the better eye; age, $M = 45.9$, $SD = 13.8$, range = 22–73) used the three Web sites. The sample of visually impaired users included 38 (69%) users who were blind and 17 (31%) users with impaired eyesight. All of the blind users employed screen-reading software or a combination of a screen reader with a braille keyboard (65%). All of the 17 users with impaired eyesight used assistive tools. One user used only a screen reader, nine users used only a screen magnifier, three users used a screen reader in combination with a screen magnifier, and four users reported that they used an assistive tool without specifying it. Each participant completed five information search tasks on one of three Web sites (NA, A, or AA) and provided subjective ratings afterward. The tasks were similar to those used in the present study (see Table 2). The subjective ratings had two aims: to determine whether the created Web sites were reasonable and to verify whether the accessibility manipulations benefited people with disabilities (i.e., visual impairments).

The first rating was intended to validate whether the manipulations were reasonable and comparable to Web sites found on the Internet (“How was the overall quality of the Web site compared to Web sites you usually use?” Possible answers were *much worse* [1], *worse* [2], *equal* [3], *better* [4], and *much better* [5]). Since we downgraded the Web sites from Level AA to A and NA, the aim of this question was to check whether the Web sites were not unrealistically downgraded. Having reasonable manipulations would thus imply that people with or without disabilities would rate the A and NA Web sites similar in quality to typical Web sites found on

the Internet. The results confirmed that the confidence intervals [CIs] for both Web sites included a score of 3, which indicated that the NA and A Web sites did not differ significantly from the mean position of the scales (i.e., 3). Hence, participants perceived these Web sites to be similar in quality to Web sites they usually use (NA Web site, $M = 2.71$, $SD = 0.87$, $CI = [2.40, 3.01]$; A Web site, $M = 3.03$, $SD = 0.77$, $CI = [2.77, 3.29]$). This finding was supported by a separate analysis for the users with and without disabilities. Ratings of users with visual impairments did not differ significantly between Levels NA and A ($t = 0.85$, $df = 33$, $p > .05$; NA Web site, $M = 2.88$, $SD = 0.99$; A Web site, $M = 3.17$, $SD = 0.99$). Likewise, the ratings of non-disabled users did not differ significantly between Levels NA and A ($t = 1.74$, $df = 27$, $p > .05$; NA Web site, $M = 2.53$, $SD = 0.72$; A Web site, $M = 2.89$, $SD = 0.47$).

The second item was designed to examine the effect of the accessibility manipulations on users with visual impairments. Therefore, participants with visual impairments additionally rated the usability of the Web site by answering the question, “How usable was the Web site overall?” The response scale ranged from 1 (*not at all*) to 5 (*very*). Having valid accessibility manipulations implies that people with visual impairments give the lowest usability rating to the NA Web site, followed by the A Web site, and the best rating to the AA Web site. The results clearly confirmed this pattern with a highly significant effect, $F(2, 52) = 13.15$, $p < .001$ (NA Web site, $M = 3.24$, $SD = 0.90$; A Web site, $M = 4.00$, $SD = 0.84$; AA Web site, $M = 4.60$, $SD = 0.68$). Post hoc analyses with Bonferroni corrections revealed a significant difference between Level NA and Level A ($p = .02$) and between Level NA and Level AA ($p < .001$). Levels A and AA did not differ significantly

($p > .05$), which was to be expected because the majority of participants were blind. Most of the criteria for Level AA are intended to support users with impaired eyesight rather than blind people, whereas most Level A criteria are intended to support blind users (cf. Appendix and Cooper et al., 2014). The manipulation check thus confirmed that the Web sites were suitable and valid for manipulating the level of accessibility in a realistic manner.

Measures

Performance measures. Two measures were taken to assess performance: (a) task completion time, defined as the time (in seconds) used to complete a given task and (b) task completion rate (percentage of tasks successfully completed).

Subjective measures. Five established questionnaires were used to take the subjective measures. For all of them, German-language versions of the questionnaires were employed: (a) Perceived usability was measured by the Website Analysis and MeasureMent Inventory (WAMMI; Kirakowski, Claridge, & Whitehand, 1998); (b) perceived aesthetics was assessed by the Visual Aesthetics of Websites Inventory (VisAWI; Thiel-sch & Moshagen, 2011). (c) To measure perceived trustworthiness of the Web site, a five-item subscale of the Scale for Online Users' Trust was used (SCOUT; Bär, 2014). (d) Positive and negative affect was assessed by the Positive and Negative Affect Schedule (PANAS; Watson, Clark, & Tellegen, 1988). (e) Subjective workload was measured by the NASA Task Load Index (NASA-TLX; Hart & Staveland, 1988).

Procedure

The experiment took place in a usability lab at the University of Fribourg. Before beginning the experiment, the landing page of the Web site was briefly shown to participants to check that they did not know the Web site (none of the participants had seen it before). Afterward, five tasks had to be completed (see Table 2). If a task was not completed after 4 min, the experimenter helped the participant solve the task. In such a case, task completion time was set to a default value of 4 min, and task completion was scored as being unsuccessful. Participants began

the experiment by filling in a questionnaire about their current positive and negative affect (PANAS). While participants were using the Web sites, screen-recording software was used to assess task completion time and completion rate. After completing the tasks, each participant completed the PANAS again, followed by further questionnaires (i.e., WAMMI, VisAWI, SCOUT, NASA-TLX, and a demographic questionnaire).

Materials

Testing was conducted on a MacBook Pro (13 inches, Intel HD Graphics 4000) with an external mouse. The Web sites were navigated with the browser Mozilla Firefox 26.0.

Data Analysis

We conducted a one-factorial analysis with level of accessibility as an independent variable. Post hoc tests with Bonferroni adjustment were used to determine differences between the three experimental conditions (low accessibility, medium accessibility, high accessibility).

For measuring changes in positive and negative affect, a baseline measurement by the PANAS was taken. This baseline measurement was then used as a covariate for analyzing positive and negative affect after Web site use.

An outlier analysis was carried out for each dependent variable, employing the median absolute deviation method (Leys, Ley, Klein, Bernard, & Licata, 2013). This method is robust and especially recommended for experiments with small to medium sample sizes. A conservative threshold of 3 was chosen (Miller, 1991). Four outliers (6.6%) were detected for perceived usability and one (1.6%) for trustworthiness. They were not included in the respective data analysis.

Due to technical problems, a video file was lost. Therefore, the performance data of one participant were not included in the analysis.

RESULTS

Performance Measures

Task completion rate. Table 3 shows that there were differences in completion rates as a function of WCAG 2.0 conformance. The effect was significant, $F(2, 57) = 7.38, p = .001$, partial $\eta^2 = .21$.

TABLE 3: Performance Measures as a Function of WCAG 2.0 Conformance Levels: Means, Standard Deviations, and Results of Post Hoc Tests

Measure	NA Web Site M (SD)	A Web Site M (SD)	AA Web Site M (SD)
Task completion rate in percentage**	65.0 (15.7)	68.6 (16.5)	82.1 (12.2)
Confidence interval	[58.3, 71.6]	[61.9, 75.2]	[75.5, 88.8]
Task completion time in seconds*	136.8 (21.9)	125.95 (27.9)	116.9 (18.3)
Confidence interval	[126.5, 147.2]	[115.6, 136.3]	[106.6, 127.2]

Note. WCAG = Web Content Accessibility Guidelines; NA = very low conformance; A = low conformance; AA = high conformance.

* $p < .05$ (two tailed). ** $p < .01$ (two tailed).

As predicted, completion rates were highest for the AA Web site. Post hoc tests revealed significant differences between conditions AA and A ($p = .016, r = .43$) as well as between AA and NA ($p = .004, r = .52$). There was no significant difference between A and NA.

Task completion time. As expected, task completion time decreased with higher WCAG 2.0 conformance (see Table 3). The effect of WCAG 2.0 conformance on task completion time was significant, $F(2, 57) = 3.74, p = .03$, partial $\eta^2 = .11$. Post hoc tests showed a significant difference between Level AA and Level NA ($p = .03, r = -.50$). There was no further significant difference in pairwise comparisons.

Subjective Measures

Usability. In line with our hypothesis, Table 4 shows that ratings for perceived usability were highest for the AA Web site. The effect on perceived usability was significant, $F(2, 54) = 5.18, p = .009$, partial $\eta^2 = .16$. According to post hoc tests, Level AA differed significantly from Level A ($p = .015, r = .41$) and also from Level NA ($p = .018, r = .36$). The difference between Levels A and NA was not significant.

Aesthetics. WCAG 2.0 conformance showed a significant effect on perceived aesthetics, $F(2, 58) = 4.23, p = .019$, partial $\eta^2 = .13$. Again, highest ratings were given for Level AA. As presented in Table 4, post hoc tests showed that AA was significantly different from A ($p = .023, r = .47$). There were no further significant pairwise comparisons.

Trustworthiness. The effect of WCAG 2.0 conformance on perceived trustworthiness was also significant, $F(2, 57) = 3.47, p = .038$, partial $\eta^2 = .11$. Participants gave highest ratings for Level AA, whereas ratings in conditions A and NA were similar (see Table 4). However, post hoc tests did not reveal any significant pairwise comparison.

Affect. The data for both measures of affect are shown in Table 4. Accessibility did not show an effect on positive affect, $F(2, 57) = 2.25, p = .115$, partial $\eta^2 = .07$. The same applied to negative affect, $F(2, 57) = 0.371, p = .691$, partial $\eta^2 = .01$.

Workload. As predicted, ratings of perceived workload were lowest for the AA Web site (see Table 4). The effect on perceived workload was significant, $F(2, 58) = 6.23, p = .004$, partial $\eta^2 = .18$. As shown in Table 4, post hoc tests revealed a significant difference between conditions AA and NA ($p = .005, r = -.45$) and between the levels AA and A ($p = .024, r = -.38$) but not between the other WCAG 2.0 levels (see Table 4). An analysis of the subscales of the workload questionnaire revealed that accessibility affected the dimensions mental demands, $F(2, 58) = 7.25, p = .002$, partial $\eta^2 = .2$, and effort, $F(2, 58) = 7.6, p = .001$, partial $\eta^2 = .2$, but none of the other dimensions (see Table 5).

DISCUSSION

We aimed to investigate consequences of implementing accessibility guidelines for non-disabled users. Employing Web sites based on the current accessibility guidelines (i.e., WCAG 2.0), a standardized experimental approach was

TABLE 4: Subjective Measures as a Function of WCAG 2.0 Conformance Levels: Means, Standard Deviations, and Results of Post Hoc Tests

Measure	NA Web Site M (SD)	A Web Site M (SD)	AA Web Site M (SD)
Usability (LS: 1–5)**	2.8 (0.5)	2.8 (0.3)	3.3 (0.7)
Confidence interval	[2.6, 3.1]	[2.5, 3.1]	[3.1, 3.6]
Aesthetics (LS: 1–7)*	4.6 (1.1)	4.3 (0.8)	5.1 (0.8)
Confidence interval	[4.2, 5.0]	[3.9, 4.7]	[4.7, 5.6]
Trustworthiness (LS: 1–5)*	3.7 (0.8)	3.7 (0.5)	4.1 (0.6)
Confidence interval	[3.4, 3.9]	[3.5, 4.0]	[3.9, 4.4]
Workload (LS: 1–20)*	10.9 (2.1)	10.5 (2.3)	8.4 (2.9)
Confidence interval	[9.9, 12.0]	[9.4, 11.6]	[7.3, 9.5]
Positive affect (LS: 1–5) ^{ns}	2.6 (0.6)	2.5 (0.8)	2.6 (0.6)
Negative affect (LS: 1–5) ^{ns}	1.3 (0.3)	1.5 (0.6)	1.3 (0.4)

Note. WCAG = Web Content Accessibility Guidelines; NA = very low conformance; A = low conformance; AA = high conformance; LS = Likert scale.

* $p < .05$ (two tailed). ** $p < .01$ (two tailed).

used to test nondisabled users, using a larger sample of nondisabled users than previous studies. Measures of performance (i.e., task completion time and task completion rate) and user ratings (i.e., perceived usability, aesthetics, workload, trustworthiness, and affect) were taken. A Web site was modified to meet the three WCAG 2.0 conformance levels (i.e., NA, A, and AA). It was expected that increasing WCAG 2.0 conformance would benefit user performance and user evaluations. The AA Web site showed advantages over the two other Web sites with regard to performance and subjective evaluations. No differences were found between NA and A.

Performance

The present results support the assumption that a Web site’s higher WCAG 2.0 conformance would lead to higher task completion rates and lower task completion time. Participants using the AA Web site were more successful in solving tasks and were faster in doing so than participants who used Web sites A or NA. The results are in line with previous work that also showed an increase in performance of nondisabled users with higher accessibility (Disability Rights Commission, 2004; Petrie & Kheir, 2007) and no difference between NA and A (Pascual et al., 2014).

Subjective Measures

It was hypothesized that participants’ evaluations would be more positive for Web sites with higher WCAG 2.0 conformance than for Web sites with lower conformance. As predicted, participants using the AA Web site gave higher ratings of usability, aesthetics, and trustworthiness and lower ratings in workload than participants in the other conditions. For most of the dependent variables, the post hoc analysis revealed that there was a difference between Levels AA and NA but not between the other conditions. These results correspond to the findings of previous work showing that higher WCAG 1.0 conformance is also positively associated with perceived usability of nondisabled users (Huber & Vitouch, 2008). Furthermore, the present results support Petrie et al.’s (2004) view that accessibility can influence visual design. The finding that accessibility is positively related to aesthetics found for WCAG 1.0 (Mbipom, 2009) is now replicated for WCAG 2.0.

A closer look into the relationship between accessibility and workload revealed that only the dimensions mental demands and effort were affected by the accessibility manipulation. The other dimensions of the NASA-TLX—physical demands, temporal demands,

TABLE 5: NASA Task Load Index Subscales as a Function of WCAG 2.0 Conformance Levels: Means, Standard Deviations, and Results of Post Hoc Tests

Measure	NA Web Site M (SD)	A Web Site M (SD)	AA Web Site M (SD)
Mental demands (LS: 1–20)**	13.5 (2.1)	11.9 (3.0)	9.6 (4.6)
Confidence interval	[12.1, 15.0]	[10.4, 13.4]	[8.1, 11.1]
Effort (LS: 1–20)**	12.4 (2.3)	11.6 (3.4)	8.2 (4.8)
Confidence interval	[10.8, 14.0]	[9.9, 13.2]	[6.6, 9.8]
Physical demands (LS: 1–20) ^{ns}	7.1 (5.1)	5.2 (5.2)	4.9 (3.4)
Temporal demands (LS: 1–20) ^{ns}	10.0 (3.5)	10.9 (4.3)	9.15 (4.4)
Performance (LS: 1–20) ^{ns}	8.9 (4.1)	9.0 (4.3)	11.4 (5.0)
Frustration (LS: 1–20) ^{ns}	10.5 (4.7)	11.6 (5.3)	9.0 (6.0)

Note. WCAG = Web Content Accessibility Guidelines; NA = very low conformance; A = low conformance; AA = high conformance; LS = Likert scale.
* $p < .05$ (two tailed). ** $p < .01$ (two tailed).

own performance, and frustration—were not significantly influenced by different levels of accessibility. According to the WCAG 2.0, the aim of the manipulated recommendations is to increase the understandability (e.g., Criterion 1.3.1, “info and relationship,” or Criterion 1.4.8, “visual presentation”), operability (e.g., Criterion 2.4.4, “link purpose,” or Criterion 2.4.10, “section headings”), and understandability (e.g., Criterion 3.4.4, “consistent identification,” or Criterion 3.3.1, “error identification”) (see the appendix for descriptions of the criteria; Cooper et al., 2014). It is thus in our opinion plausible that these recommendations reduce mental demands and effort by providing a clearer structure of the Web site (e.g., by providing section headings), increasing predictability (e.g., due to purposeful link texts), or providing higher distinguishability of elements (e.g., by a consistent usage of styles for Web site elements). This assumption is supported by previous research that showed that reducing complexity in Web sites is associated with lower perceived workload (e.g., Schmutz, Heinz, Métrailler, & Opwis, 2009).

Physical and temporal demands were not affected because browsing a Web site is usually not physically demanding, and there was no obvious time limit for solving the tasks. Participants might thus not have experienced time pressure. An effect on frustration might not have

occurred because participants did not receive a feedback about their performance. Additionally, it was emphasized that the experiment focused on the evaluation of the Web site and did not examine the performance of participants, which may also explain the lack of an effect on perceived performance. Since, to our knowledge, there is no other study focusing on the relationship of accessibility and workload, these associations may need to be investigated further.

Similarly, no study has related accessibility to perceived trustworthiness of nondisabled users. However, an earlier study showed that well-structured content (e.g., possible sequence of clicks and paths on Web sites) is an important antecedent of trust in Web sites (Bart, Shankar, Sultan, & Urban, 2005). This finding might offer an explanation for the effects of accessibility on trust, because in the presented study, the manipulated accessibility recommendations may have also influenced the structure of the Web site in a similar way (e.g., by providing section headings, meaningful labels, and purposeful link texts).

Relevant Success Criteria for Nondisabled Users

An interesting finding was that effects on nondisabled users occurred only when changing the Web site from NA to AA but not from NA to A or from A to AA. We thus assume that the combination of changes on Level A and

Level AA is responsible for positive effects on nondisabled users and that there is not a single dominant criterion (e.g., link text) that caused the effect. In case of a single dominant criterion, effects would have occurred either between NA and A or between A and AA. Nevertheless, we assume that some criteria were more important than others (e.g., users rarely used tabbing, which suggests that changing tabbing order was not relevant for the present results). The remaining criteria may jointly benefit nondisabled users by providing support in reading (i.e., contrast and text alignment), completing a form (i.e., form labeling and error identification), and navigating by providing clear structure (i.e., meaningful headings and link texts). User comments in postexperimental interviews support the assumption of a combined effect of success criteria because participants mainly gave general comments, such as “good structure” or “very clear website” for the AA Web site and “sometimes unclear structure” or “a rather complex Web site” for the NA Web site. Nobody mentioned a specific issue, such as unclear link text or low contrast.

Common Recommendations for Users With and Without Disabilities

It is important to note that many accessibility requirements found in WCAG 2.0 are also recommended in guidelines for designing user-friendly Web sites or interfaces for nondisabled people (e.g., Farkas & Farkas, 2000; Nielsen, Tahir, & Tahir, 2002; Shneiderman & Ben, 2003; Spool, 1999; Spyridakis, 2000; Williams, 2000). Examples for such recommendations are to use precise link texts, use headings to structure the content, use consistent design, and use left-aligned text. The overlap between recommendations for accessibility and recommendations for good Web site design for nondisabled users strengthens our assumption of beneficial effects of accessibility on nondisabled users and may partly explain the present results.

Limitation and Future Research

The present study has some limitations. First, to complement previous research, we chose a controlled experimental approach by

manipulating a single Web site according to different accessibility levels with a view to eliminating Web site characteristics that are not related to accessibility guidelines (e.g., differences in written content, images, or type of Web site). The downside of this approach is that the NA and A Web sites used do not actually exist. However, we tried to obtain reasonable representations of Levels NA and A by considering the literature on typical accessibility issues, and we used existing Web sites as a reference for our manipulations. Furthermore, the validation study revealed that the manipulations were reasonable for users with and without disabilities. We think that studies emphasizing high experimental control should complement (though not replace) work focusing on making comparisons between different existing Web sites. Therefore, authors of future research should pursue both paths. Second, the type of Web site used in the present study (an existing local government Web site) is not necessarily representative of the wide range of Web sites found on the Web. For example, the Web site did not contain any multimedia content (e.g., video content) or interaction elements (e.g., drag-and-drop or captcha). Manipulating the accessibility of such Web site features may result in different effects on nondisabled users and needs to be addressed in future research. Nevertheless, the characteristics of the Web site used in this study are comparable to many types of Web sites, including Web sites of industry, educational institutions, blogs, and news.

Third, user ratings may not be sufficient to gauge complex concepts, such as affect or aesthetics. Authors of future research could take this into account by using subjective measures together with objective measures (e.g., physiological measures for assessing emotional reactions). Fourth, accessibility is a complex concept with different components, such as users, developers, and content (Chisholm & Henry, 2005). Therefore, authors of future research should compare effects of accessibility guidelines on both nondisabled users and people with disabilities rather than focusing on nondisabled users alone. This comparison will allow us to gain a deeper understanding of the relation between accessibility and nondisabled users. Fifth, since the sample comprised young students, the

generalizability of present results to a more heterogeneous population may be limited. Authors of future research may investigate effects of accessibility on nondisabled users from a wider range of age and educational level. However, we would expect even stronger effects of accessibility on older or less educated samples because they might be less experienced in using Web sites than young students and would benefit more from supportive Web site characteristics.

Implications for Practitioners

The present study has important implications for practice. First, WCAG 2.0 should be considered not only as an aid for designing Web sites according to the needs of users with disabilities but also as a helpful tool for designing more usable Web sites for nondisabled users. This different framing may motivate practitioners to apply these guidelines more often (because of the benefits to nondisabled users) while alleviating the financial concerns of practitioners about Web site accessibility. As an implication for the guidelines, positive effects for users without disabilities should be mentioned explicitly as well as the fact that Level AA is of particular importance for such users. Second, the consistent pattern of beneficial effects of Level AA compared to NA is highly relevant for practitioners. Currently, most of the Web sites conform to Level NA (e.g., Gonçalves et al., 2013; Nurmela et al., 2013), which shows that there is much room for improvement. Practitioners should aim for an upgrade from Level NA to AA rather than A (since the latter would not provide noticeable benefits to nondisabled users).

The effect sizes between the conditions NA and AA were consistently medium to large. For instance, the AA Web site led to a mean decrease in task completion time of about 20 s (i.e., 15%) and increased the task completion rate by about 17%. Third, the 10 success criteria that were changed from Level NA to AA are rather easy to implement (e.g., meaningful link text, sufficient contrast, text alignment). These “easy-to-be-changed” criteria may help practitioners improve Web sites or design new ones, following WCAG 2.0, by offering a positive cost-benefit trade-off.

CONCLUSION

The present work demonstrated that implementing accessibility guidelines can provide several benefits for nondisabled users. To achieve these benefits, high conformance (i.e., Level AA) to current guidelines (i.e., WCAG 2.0) is necessary. Overall, the research field of accessibility still seems to be virgin territory despite its important impact on society and the considerable number of users that are affected. Especially, effects of accessibility standards on various user groups have been hardly addressed. The present work has thus some elements of an exploratory study, which may initiate further research into this issue. This research is important because further knowledge might lead to increasing awareness and acceptance of accessibility in research and practice. We hope that our research represents a contribution to increasing the prevalence of accessible Web sites and, more generally, to the promotion of equality.

APPENDIX

Overview of Web Site Characteristics Manipulated

Success Criterion (WCAG 2.0 Level)	AA Web Site	A Web Site	NA Web Site	Comments, References to the Document Understanding WCAG 2.0
1.1.1 Nontext content (A)	Every image on the Web site has an appropriate text alternative	Every image on the Web site has an appropriate text alternative	For every image, the text alternative image was used	The manipulation is based on common failures F30 and F39 (Manipulated for further studies; did not affect nondisabled users)
1.3.1 Info and relationships (A)	Required fields in the contact form were labeled with bold text and with an asterisk with the text alternative required	Required fields in the contact form were labeled with bold text and with an asterisk with the text alternative required	**Required fields were labeled only with bold text, without text alternative required	The manipulation represents a violation of the sufficient technique G117
1.4.3 Contrast (minimum) (AA)	The contrast between headings and background was 4.5:1 (#007FAF #FFFFFF) The contrast between text and background was 21.0:1 (#FFFFFF #000000)	*The contrast between headings and background was 3.9:1 (#007FEF #FFFFFF) The contrast between text and background was 4.0:1 (#FFFFFF #7F7F7F)	*The contrast between headings and background was 3.9:1 (#007FEF #FFFFFF) The contrast between text and background was 4.0:1 (#FFFFFF #7F7F7F)	The manipulation represents a violation of the sufficient technique G18 The chosen contrasts seem to be realistic since the screening revealed that there are plenty of Web sites containing contrasts about 3.0:1 or lower
1.4.4 Resize text (AA)	Text can be resized without assistive technology up to 200% without loss of content or functionality	Resizing text to 200% caused text passages to be truncated or obscured	Resizing text to 200% caused text passages to be truncated or obscured	The manipulation is based on common failure F69 (Manipulated for further studies; did not affect nondisabled users because nobody resized text)
1.4.8 Visual presentation (AAA)	Text blocks had a maximum width of 80 characters and were left aligned	*Text blocks had a maximum width of 90 characters and were justified	*Text blocks had a maximum width of 90 characters and were justified	The manipulation is based on common failure F88 as well as a violation of sufficient technique C20

(continued)

APPENDIX (continued)

Success Criterion (WCAG 2.0 Level)	AA Web Site	A Web Site	NA Web Site	Comments, References to the Document <i>Under- standing WCAG 2.0</i>
2.4.3 Focus order (A)	Focusable components receive focus in an order that preserves meaning and operability	Focusable components receive focus in an order that preserves meaning and operability	**Some fields in the form did not receive focus in a typical order via tabbing (i.e., skips between fields in different sections of the form. Focus moved from the name field to a checkbox above, then to the street address)	The manipulation is based on common failure F44 as well as Example 5 of understanding Success Criterion 2.4.3
2.4.4 Link purpose (in context) (A)	Links for sending an e-mail to a certain person were presented as mail address (e.g., john. smith@example. com—the purpose can be determined from the link text only)	*Links for sending an e-mail to a certain person were labeled <i>contact</i> within the same paragraph as the description of the respective person (the purpose can be determined from the link text together with its context)	**Links for sending an e-mail to a certain person were labeled <i>link</i> within the same paragraph as the description of the respective person (the purpose can not be determined with certainty from the link text together with its context)	The manipulation is based on common failure F88 as well as a violation of sufficient technique C20
2.4.6 Heading and labels (AA)	Heading and labels describe topic and purpose	*Some headings were shortened to be less descriptive (e.g., from "Information About the Town Eschen" to "General Information")	*Some headings were shortened to be less descriptive (e.g., from "Information About the Town Eschen" to "General Information")	The manipulation represents a violation of the sufficient technique G117
2.4.7 Focus visible (AA)	Keyboard focus indicator was visible	*Keyboard focus indicator was not visible	*Keyboard focus indicator was not visible	The manipulation is based on common failure F78
2.4.10 Section headings (AAA)	Section headings were used to organize the content	*Some section headings were removed	*Some section headings were removed	The manipulation represents a violation of the sufficient techniques G141 and H69

(continued)

APPENDIX (continued)

Success Criterion (WCAG 2.0 Level)	AA Web Site	A Web Site	NA Web Site	Comments, References to the Document <i>Understanding WCAG 2.0</i>
3.2.3 Consistent navigation (AA)	Navigational mechanisms occurred in the same relative order	Some navigation links were not presented in the same order on some Web pages but only in the html file (i.e., remarkable when using a screen reader) <i>The change in position was not visible due to holding the position via CSS</i>	Some navigation links were not presented in the same order on some Web pages but only in the html file (i.e., remarkable when using a screen reader) <i>The change in position was not visible due to holding the position via CSS</i>	Not remarkable without using screen-reading software <i>The manipulation is based on common failure F66</i>
3.4.4 Consistent identification (AA)	Links are designed consistently bold, in blue color, and underlined	*Links differ in design: Links were either blue and not underlined or underlined and in the same color as text; links were also not consistently bold	*Links differ in design: Links were either blue and not underlined or underlined and in the same color as text; links were also not consistently bold	The problem frequently occurred in the screening
3.3.1 Error identification (A) and 3.3.3 Error suggestions (AA)	If an input error in the form was automatically detected, the item that was in error was detected and described by the user in text (i.e., the field was marked with a red square and a textual suggestion on how to complete the field)	If an input error in the form was automatically detected, the item that was in error was detected and described by the user in text (i.e., the field was marked with a red square and a textual suggestion on how to complete the field)	**There was no error identification used in the form	The manipulation is based on a violation of sufficient technique G83

Note. WCAG = Web Content Accessibility Guidelines; NA = very low conformance; A = low conformance; AA = high conformance; CSS = Cascading Style Sheets. Modifications from Level AA to level A are highlighted with one asterisk (*). Modifications from Level A to NA are highlighted with two asterisks (**). Common failures and sufficient techniques are mentioned referring to the document *Understanding WCAG 2.0* (Cooper, Kirkpatrick, & O Connor, 2014). Since the Web sites are also used in other studies including people with visual impairments, some changes were made that clearly do not affect sighted users (e.g., changing alternative text of a picture). These changes appear in italics.

ACKNOWLEDGMENTS

We are grateful to Filippo Nugara, Andreas Uebelbacher, Catarina Koch, and Karin Waespe for their help in carrying out the research.

KEY POINTS

- Web accessibility guidelines (i.e., Web Content Accessibility Guidelines [WCAG] 2.0) may also provide benefits to nondisabled users in terms of improved performance and subjective ratings.
- Using WCAG 2.0 Level A or Level AA did not entail any detrimental effects for nondisabled users.
- Implementing high conformance to WCAG 2.0 (i.e., Level AA) is recommended to practitioners because it addresses the needs of users with and without disabilities.

REFERENCES

- Andreasen, M. S., Nielsen, H. V., Schröder, S. O., & 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. doi:10.1145/1240624.1240838
- Arrue, M., Fajardo, I., Lopez, J. M., & Vigo, M. (2007). Interdependence between technical web accessibility and usability: Its influence on web quality models. *International Journal of Web Engineering and Technology*, 3, 307–328. doi:10.1504/IJWET.2007.012059
- Badre, A. (2002). *Shaping Web usability: Interaction design in context*. Boston, MA: Addison-Wesley Professional.
- Bär, N. (2014). *Human-computer interaction and online users' trust* (Doctoral dissertation). Chemnitz University of Technology, Chemnitz, Germany. Retrieved from <http://nbn-resolving.de/urn:nbn:de:bsz:ch1-qucosa-149685>
- Bart, Y., Shankar, V., Sultan, F., & Urban, G. L. (2005). Are the drivers and role of online trust the same for all web sites and consumers? A large-scale exploratory empirical study. *Journal of Marketing*, 69, 133–152. doi:10.1509/jmkg.2005.69.4.133
- Caldwell, B., Cooper, M., Reid, L. G., & Vanderheiden, G. (2008). *Web content accessibility guidelines 2.0, World Wide Web Consortium (W3C) recommendation*. Retrieved from <http://www.w3.org/TR/2008/REC-WCAG20-20081211/>
- Chadwick-Dias, A., McNulty, M., & Tullis, T. (2003). Web usability and age: How design changes can improve performance. *ACM SIGCAPH Computers and the Physically Handicapped*, 73/74, 30–37. doi:10.1145/960201.957212
- Chisholm, W. A., & Henry, S. L. (2005). Interdependent components of Web accessibility. In *Proceedings of the 2005 International Cross-Disciplinary Workshop on Web Accessibility (W4A)* (pp. 31–37). New York, NY: ACM.
- Cooper, M., Kirkpatrick, A., & O Connor, J. (2014). *Understanding WCAG 2.0, World Wide Web Consortium (W3C) note*. Retrieved from <http://www.w3.org/TR/UNDERSTANDING-WCAG20/>
- Disability Rights Commission. (2004). *The Web: Access and inclusion for disabled people. A formal investigation*. London, UK: Stationery Office.
- Dray, S., & Siegel, D. (2004). Remote possibilities? International usability testing at a distance. *Interactions*, 11(2), 10–17. doi:10.1145/971258.971264
- Ellcessor, E. (2014). <ALT="Textbooks">: Web accessibility myths as negotiated industrial lore. *Critical Studies in Media Communication*, 31(2), 1–16. doi:10.1080/15295036.2014.919660
- Farkas, D. K., & Farkas, J. B. (2000). Guidelines for designing Web navigation. *Technical Communication*, 47, 341–358.
- Farrelly, G. (2011). Practitioner barriers to diffusion and implementation of Web accessibility. *Technology and Disability*, 23, 223–232. doi:10.3233/TAD-2011-0329
- Freire, A. P., Petrie, H., & Power, C. (2011). Empirical results from an evaluation of the accessibility of websites by dyslexic users. In *Proceedings of the Workshop on Accessible Design in the Digital World* (pp. 41–53). Lisboa, Portugal: CEUR-WS.org.
- Freire, A. P., Russo, C. M., & Fortes, R. P. M. (2008). The perception of accessibility in Web development by academy, industry and government: A survey of the Brazilian scenario. *New Review of Hypermedia and Multimedia*, 14, 149–175. doi:10.1080/13614560802624241
- Gonçalves, R., Martins, J., Pereira, J., Oliveira, M. A., & Ferreira, J. J. P. (2013). Enterprise Web accessibility levels amongst the Forbes 250: Where art thou o virtuous leader? *Journal of Business Ethics*, 113, 363–375. doi:10.1007/s10551-012-1309-3
- Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. *Advances in Psychology*, 52, 139–183. doi:10.1016/S0166-4115(08)62386-9
- Henry, S. L. (2006). Understanding Web accessibility. In J. Thatcher, R. M. Burks, C. Heilmann, S. L. Henry, A. Kirkpatrick, P. H. Lauke . . . C. D. Waddell (Eds.), *Web accessibility: Web standards and regulatory compliance* (pp. 1–51). New York, NY: Apress. doi:10.1007/978-1-4302-0188-5_1
- Huber, W., & Vitouch, P. (2008). Usability and accessibility on the Internet: Effects of accessible Web design on usability. In *Proceedings of the 11th International Conference, ICCHP* (pp. 482–489). Berlin, Germany: Springer. doi:10.1007/978-3-540-70540-6_69
- International Organization for Standardization. (2012). *ISO 40500:2012. W3C Web content accessibility guidelines (WCAG) 2.0*. Geneva, Switzerland: Author.
- Johnson, R., & Kent, S. (2007). Designing universal access: Web-applications for the elderly and disabled. *Cognition, Technology & Work*, 9, 209–218.
- Kirakowski, J., Claridge, N., & Whitehand, R. (1998, June). *Human centered measures of success in Web site design*. Paper presented at the Fourth Conference on Human Factors and the Web, Basking Ridge, NJ. Retrieved from <http://research.microsoft.com/en-us/um/people/marycz/hfweb98/kirakowski/>
- Kurniawan, S., & Zaphiris, P. (2005). Research-derived Web design guidelines for older people. In *Proceedings of the 7th International ACM SIGACCESS Conference on Computers and Accessibility* (pp. 129–135). New York, NY: ACM.
- Lawson, B. (2006). Introduction. In J. Thatcher, R. M. Burks, C. Heilmann, S. L. Henry, A. Kirkpatrick, P. H. Lauke . . . C. D. Waddell (Eds.), *Web accessibility: Web standards and regulatory compliance* (pp. xxv–xlvi). New York, NY: Springer Verlag.
- Lazar, J., Dudley-Sponaule, A., & Greenidge, K. D. (2004). Improving Web accessibility: A study of Webmaster perceptions. *Computers in Human Behavior*, 20, 269–288. doi:10.1016/j.chb.2003.10.018

- Leys, C., Ley, C., Klein, O., Bernard, P., & Licata, L. (2013). Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median. *Journal of Experimental Social Psychology*, 49, 764–766. doi:10.1016/j.jesp.2013.03.013
- Loiacono, E. T., & Djasasbi, S. (2013). Corporate Website accessibility: Does legislation matter? *Universal Access in the Information Society*, 12, 115–124. doi:10.1007/s10209-011-0269-1
- Mbipom, G. (2009). Good visual aesthetics equals good Web accessibility. *ACM SIGACCESS Accessibility and Computing*, 93, 75–83. doi:10.1145/1531930.1531939
- Mead, S. E., Lamson, N., & Rogers, W. A. (2002). Human factors guidelines for Web site usability: Health-oriented Web sites for older adults. In R. W. Morrell (Ed.), *Older adults, health information, and the World Wide Web* (pp. 89–107). Hillsdale, NJ: Lawrence Erlbaum.
- Miller, J. (1991). Short report: Reaction time analysis with outlier exclusion. Bias varies with sample size. *Quarterly Journal of Experimental Psychology*, 43, 907–912. doi:10.1080/14640749108400962
- Mlynarczyk, G. (2012, July 26). 7 Web accessibility myths [Web log post]. Retrieved from <http://www.nomensa.com/blog/2012/7-web-accessibility-myths-2/>
- Murphy, S. K. (2013). WCAG 2.0: Emerging international standard for Web accessibility and video [Web log post]. Retrieved from <http://www.3playmedia.com/2013/11/25/wcag-2-0-emerging-international-standard-web-accessibility-video-captions/>
- Nayak, L., Priest, L., Stuart-Hamilton, I., & White, A. (2006). Website design attributes for retrieving health information by older adults: An application of architectural criteria. *Universal Access in the Information Society*, 5, 170–179.
- Nielsen, J., Tahir, M., & Tahir, M. (2002). *Homepage usability: 50 Websites deconstructed*. Indianapolis, IN: New Riders.
- Nurmela, K., Pirhonen, A., & Salminen, A. (2013). Accessibility of public Web services: A distant dream? In P. Kotzé, G. Marsden, G. Lindgaard, J. Wesson, & M. Winckler (Eds.), *Human-Computer Interaction-INTERACT 2013* (pp. 566–578). Berlin, Germany: Springer. doi:10.1007/978-3-642-40483-2_40
- Pak, R., & Price, M. M. (2008). Designing an information search interface for younger and older adults. *Human Factors*, 50, 614–628.
- Pascual, A., Ribera, M., Granollers, T., & Coiduras, J. L. (2014). Impact of accessibility barriers on the mood of blind, low-vision and sighted users. *Procedia Computer Science*, 27, 431–440. doi:10.1016/j.procs.2014.02.047
- Petrie, H., Hamilton, F., & King, N. (2004). Tension, what tension? Website accessibility and visual design. In *Proceedings of the 2004 International Cross-Disciplinary Workshop on Web Accessibility (W4A)* (pp. 13–18). New York, NY: ACM. doi:10.1145/990657.990660
- Petrie, H., & Kheir, O. (2007). The relationship between accessibility and usability of Websites. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 397–406). New York, NY: ACM. doi:10.1145/1240624.1240688
- Power, C., Freire, A., Petrie, H., & Swallow, D. (2012). Guidelines are only half of the story: Accessibility problems encountered by blind users on the Web. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 433–442). New York, NY: ACM. doi:10.1145/2207676.2207736
- 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-Disciplinary Conference on Web Accessibility (W4A)* (pp. 109–115). New York, NY: ACM. doi:10.1145/1368044.1368069
- Reid, L. G., & Snow-Weaver, A. (2009). WCAG 2.0 for designers: Beyond screen readers and captions. In C. Stephanidis (Ed.), *Universal access in human-computer interaction: Applications and services* (pp. 674–682). Berlin, Germany: Springer. doi:10.1007/978-3-642-02713-0_71
- Rello, L., Kanvinde, G., & Baeza-Yates, R. (2012). Layout guidelines for web text and a web service to improve accessibility for dyslexics. In *Proceedings of the International Cross-Disciplinary Conference on Web Accessibility* (Article 36). New York, NY: ACM.
- Rogers, M. (2015). Government accessibility standards and WCAG 2 [Web log post]. Retrieved from <http://www.powermapper.com/blog/government-accessibility-standards/>
- Rogers, W. A., & Fisk, A. D. (Eds.). (2001). *Human factors interventions for the health care of older adults*. Boca Raton, FL: CRC Press.
- Rømen, D., & Svanæs, D. (2012). Validating WCAG Versions 1.0 and 2.0 through usability testing with disabled users. *Universal Access in the Information Society*, 11, 375–385. doi:10.1007/s10209-011-0259-3
- Ruth-Janneck, D. (2011). Experienced barriers in Web applications and their comparison to the WCAG guidelines. In A. Holzinger & K. Simonic (Eds.), *Information quality in e-health* (pp. 283–300). Berlin, Germany: Springer. doi:10.1007/978-3-642-25364-5_21
- Sayago, S., Camacho, L., & Blat, J. (2009). Evaluation of techniques defined in WCAG 2.0 with older people. In *Proceedings of the 2009 International Cross-Disciplinary Conference on Web Accessibility (W4A)* (pp. 79–82). New York, NY: ACM. doi:10.1145/1535654.1535673
- Schmutz, P., Heinz, S., Métrailler, Y., & Opwis, K. (2009). Cognitive load in e-commerce applications: Measurement and effects on user satisfaction. *Advances in Human-Computer Interaction*, 2009, 3. doi:10.1155/2009/121494
- Shneiderman, B., & Ben, S. (2003). *Designing the user interface*. Delhi, India: Pearson Education India.
- Spool, J. M. (1999). *Web site usability: A designer's guide*. Burlington, MA: Morgan Kaufmann.
- Spyridakis, J. H. (2000). Guidelines for authoring comprehensible Web pages and evaluating their success. *Technical Communication*, 47, 359–382.
- Sullivan, T., & Matson, R. (2000). Barriers to use: Usability and content accessibility on the Web's most popular sites. In *Proceedings on the 2000 Conference on Universal Usability* (pp. 139–144). New York, NY: ACM. doi:10.1145/355460.355549
- Thielsch, M. T., & Moshagen, M. (2011). Erfassung visueller Ästhetik mit dem VisAWI [Measuring visual aesthetics with the VisAWI]. *Usability Professionals*, 9, 260–265.
- Vu, K. P. L., & Proctor, R. W. (Eds.). (2011). *Handbook of human factors in Web design*. Boca Raton, FL: CRC Press.
- Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology*, 54, 1063–1070. doi:10.1037/0022-3514.54.6.1063
- Westerman, S. J., Davies, D. R., Glendon, A. I., Stammers, R. B., & Matthews, G. (1995). Age and cognitive ability as predictors of computerized information retrieval. *Behaviour & Information Technology*, 14, 313–326. doi:10.1080/01449299508914650

- Williams, T. R. (2000). Guidelines for designing and evaluating the display of information on the Web. *Technical Communication*, 47, 383–396.
- World Health Organization. (2011). *World report on disability*. Geneva, Switzerland: WHO Press.
- Yesilada, Y., Brajnik, G., Vigo, M., & Harper, S. (2012). Understanding Web accessibility and its drivers. In *Proceedings of the International Cross-Disciplinary Conference on Web Accessibility* (Article 19). New York, NY: ACM. doi:10.1145/2207016.2207027
- Yesilada, Y., Brajnik, G., Vigo, M., & Harper, S. (2013). Exploring perceptions of Web accessibility: A survey approach. *Behaviour & Information Technology*, 34, 119–134. doi:10.1080/0144929X.2013.848238

Sven Schmutz is a PhD student at the University of Fribourg, Switzerland. He received his MS in work and organizational psychology from the University of Bern, Switzerland.

Andreas Sonderegger is a lecturer at the Department of Psychology, University of Fribourg, Switzerland and UX Analyst at EPFL+Ecal-Lab, EPFL Lausanne, Switzerland. He received his PhD in psychology from the University of Fribourg, Switzerland in 2010.

Juergen Sauer is Professor of Cognitive Ergonomics at the Department of Psychology, University of Fribourg, Switzerland. He received an MSc in occupational psychology from the University of Sheffield, UK in 1990 and a PhD in psychology from the University of Hull, UK in 1997.

Date received: May 29, 2015

Date accepted: February 14, 2016