Applied Ergonomics 74 (2019) 97–106

Contents lists available at ScienceDirect

Applied Ergonomics

journal homepage: www.elsevier.com/locate/apergo

Easy-to-read language in disability-friendly web sites: Effects on nondisabled users

Sven Schmutz*, Andreas Sonderegger, Juergen Sauer

Department of Psychology, University of Fribourg, Fribourg, Switzerland

ARTICLE INFO

Keywords:
Web accessibility
Easy-to-read language
Nondisabled users

ABSTRACT

This article examines the influence of using easy-to-read language (as an important web accessibility criterion to support users with disabilities) on nondisabled users. A web site usability test was conducted with 128 non-disabled participants. Each participant evaluated one of two web site versions: conventional language vs. easy-to-read language. Measures of performance (e.g., recognition of content and reading time) and subjective user reactions (e.g., perceived usability and text liking) were taken. The results showed that easy-to-read language (compared to conventional language) led to improvements for some performance measures (e.g., recognition of content) but also to increased reading time, decreased text liking and reduced intention to revisit the web site. This article concludes that when implementing easy-to-read language in web sites, adverse as well as beneficial effects on nondisabled users need to be considered.

1. Introduction

1.1. Web accessibility

The accessibility of web sites for people with disabilities is an issue of growing importance in the field of human computer interaction (Jacko, 2012; Vu and Proctor, 2011). In many cases, people with impairments (e.g., visual, motor, hearing or cognitive) face barriers when using web sites because the design does not consider their needs (Henry, 2006; Thatcher et al., 2006). This is of concern because a considerable proportion of the population has some kind of impairment (i.e., about 15%; WHO, 2011). In a society heavily relying on web technologies, restricted accessibility to web-based information may cause major disadvantages. This may include fewer opportunities in education or the labour market, which, in turn, may cause costs because the people concerned are less independent and have to be financially supported by society. Therefore, the approach of making the 'web accessible' aims to reduce barriers on web sites for people with disabilities (e.g., Henry, 2006; Thatcher et al., 2006). For example, this refers to recommendations such as having a minimum contrast between font and background to consider users with visual impairments; using captions for audio content to consider people who are deaf; or using easy-to-read language (hereafter "ETRL"), which may support users with cognitive, hearing or visual impairments (e.g. Caldwell et al., 2008; Ruth-Janneck, 2011).

1.2. Web accessibility and nondisabled users

Despite the importance of web accessibility, research revealed that, at least, 95% of web sites have not followed recommendations on accessible web site design (Gonçalves et al., 2013; Nurmela et al., 2013). An important reason for the low implementation in practice may be that practitioners are concerned about negative consequences of making web sites accessible. Such negative beliefs are very common among practitioners, such as making a web site accessible results in a boring, dull or aesthetically unappealing design for nondisabled users (e.g., Ellcessor, 2014; Thatcher et al., 2006). Obviously, any possible side effects of accessible web site design for nondisabled users are of great importance for practitioners because most users are nondisabled. Although this is a well-known issue (e.g., Ellcessor, 2014; Henry, 2006; Petrie et al., 2004; Thatcher et al., 2006), only little research has begun to investigate possible effects of accessible web site design for people who are nondisabled (e.g., Pascual et al., 2014; Schmutz et al., 2016; Schmutz et al., 2017a,b). The few studies available focused on various accessible design criteria (e.g., high contrast, text alignment, tab order or heading structure) and their consequences for nondisabled users in terms of performance and subjective user reactions. The results of these studies typically indicated positive effects of accessible web site design for nondisabled users (e.g., higher task completion rates, shorter task completion times and higher ratings in perceived usability), which do not support the concerns of practitioners. While previous work already

* Corresponding author. University of Fribourg, Department of Psychology, Rue de Faucigny 2, 1700 Fribourg, Switzerland.
E-mail addresses: sven.schmutz@unifr.ch (S. Schmutz), andreas.sonderegger@unifr.ch (A. Sonderegger), juergen.sauer@unifr.ch (J. Sauer).

https://doi.org/10.1016/j.apergo.2018.08.013
Received 7 August 2017; Received in revised form 22 March 2018; Accepted 11 August 2018
Available online 20 August 2018
0003-6870/

examined the consequences for nondisabled users of certain accessible design criteria (e.g., contrast, visible focus and tabbing), very little is known about the effects of simplified language, which may be one of the most important accessibility characteristics (Ruth-Janneck, 2011).

1.3. Language complexity as accessibility criterion

A previous study (Ruth-Janneck, 2011) examined characteristics and importance of accessibility problems for users with various types of impairments (e.g., visual, hearing and cognitive impairments). The result showed overall, language complexity is the most critical barrier to access information on web sites for users with different types of impairments, such as people with visual, hearing and cognitive ones. The author concluded that simplifying text (e.g., avoiding or explaining foreign words) is an adequate intervention to tackle this issue. This finding illustrates the relevance of language complexity as accessibility criterion. While there is little research on language complexity as a web accessibility criterion, there are some practitioner guidelines with recommendations to reduce language complexity.

1.4. Language complexity in accessibility standards and guidelines

The web content accessibility guidelines 2.0 (WCAG 2.0) (Caldwell et al., 2008), as the most commonly used guidelines for web accessibility, give some recommendations for reducing language complexity. For example, the guidelines recommend providing text that does not require a reading ability that exceeds lower secondary education level (criterion 3.1.5). Furthermore, the WCAG 2.0 suggests that explanations for idioms or jargon should be provided (criterion 3.1.3) and the meaning of abbreviations should be explained (criterion 3.1.4). Although the WCAG 2.0 provides these recommendations, they are not very specific and practitioners may have problems in interpreting and implementing them adequately. There are more specific guidelines, such as the “European standard for easy-to-read“ (see http://easy-to-read.eu/?page_id=17) or the “guidelines to write clearly and simply” of WebAim (one of the leading providers of web accessibility expertise, see http://webaim.org/). Typical recommendations include: Do not use difficult words, avoid pronouns and use short sentences.

In German language, there is one particular set of guidelines for ETRL that is well established: “Regeln für Leichte Sprache des Netzwerks Leichte Sprache” (Rules for ETRL provided by the network for ETRL) (see http://www.leichtesprache.org/images/Regeln_Leichte_Sprache.pdf). The rules are similar to the European standard of easy to read but provide more specific rules on how to write easy texts. These rules comprise a total of 22 recommendations (see Appendix 1 for a complete list of the recommendations). Although there are standards providing some recommendations on how to implement ETRL, there is a lack of research investigating consequences of implementing ETRL on web sites.

1.5. Research on ETRL in web sites

While there is some basic research on effectiveness of certain ETRL recommendations (e.g., Fajardo et al., 2014; Fajardo et al., 2013; Saggion et al., 2015), there is to our knowledge only one study that empirically examined ETRL in the context of web site design (Karreman et al., 2007). In the one study investigating ETRL in the context of web site design (Karreman et al., 2007), 40 users (20 with cognitive impairments, 20 without impairments) tested two versions of a web site containing information about health services. While on one web site the information was presented in non-adapted conventional language (hereafter “ConvL”), the other web site contained text that corresponded to ETRL guidelines. These ETRL guidelines comprised recommendations on the verbal content (e.g., use short words, use words consistently, cover only one idea per sentence) and on the layout (e.g., put one sentence in one line, use a large type face, use headings and other navigational aids). Each participant completed five search tasks on one of the two web sites. The results indicated that people with cognitive disabilities understood the text better when using the ETRL-version but there was no difference in reading time for the two versions. Furthermore, the users with cognitive impairments were more satisfied when using the easy version than when using the conventional one. This suggests that the easy-to-read adaptations were supportive to their targeted user group. Regarding nondisabled users the results showed better text understanding for the ETRL version but a clear user preference for the ConvL version. This indicates that for nondisabled users, ETRL entails a trade-off between increased understanding and lower text liking. Therefore, depending on the main goal of a web site (e.g., marketing vs. information transfer), it is advisable or not to implement ETRL in web sites. This suggests that practitioners should apply ETRL with caution because it may have negative effects on nondisabled users.

The results of Karreman et al.’s (2007) study suggest that there is a difficult trade-off between making web sites more accessible to users with disabilities and meeting the needs of nondisabled people. Such conflicts are extensively discussed in the field of accessibility (e.g., Boldyreff et al., 2001; Ellcessor, 2014; Farrelly, 2011; Schmutz et al., 2016; Thatcher et al., 2006) because practitioners will not implement criteria which they believe may have adverse effects on nondisabled users, who typically represent the main user group. In line with this argument, some authors emphasise the investigation of effects of ETRL on nondisabled users (e.g., Boldyreff, et al., 2001; Karreman et al., 2007) because making a web site more accessible through ETRL should not reduce the usability for nondisabled users.

Although Karreman et al.’s study provides important first insights, some methodological issues are to be considered. Since they manipulated verbal content as well as the layout of the web sites, it is difficult to pinpoint the cause of the effect (i.e., language or layout). Furthermore, even though their study provided interesting findings in regard to text understanding and text liking, other important variables for web site design were not measured (e.g., perceived aesthetics, perceived task load, intention to revisit the web site and affect). Investigating such factors is of relevance for practitioners because knowing about possible effects of ETRL on a wide range of variables would help them to form a well-founded opinion about whether ETRL should be implemented or not.

1.6. The present study

The experiment aimed to examine the consequences of ETRL as a web accessibility criterion for nondisabled users. As independent variable, the language complexity of a web site was manipulated, which resulted in two web site versions: ConvL and ETRL. A usability test was conducted to evaluate the two web sites. Dependent variables comprised performance (e.g., detection time and free recall of content) and subjective user reactions (e.g., perceived usability and text liking). This work builds on previous research by applying a manipulation focusing only on ETRL (rather than ETRL and layout) and by investigating a wide range of dependent variables, which are relevant for the field of web site design. According to the literature review, we predicted beneficial effects of ETRL on performance but negative effects on subjective user reactions.

2. Method

2.1. Design and participants

The present experiment employed a one-factorial between-subjects design, manipulating the variable language complexity (ConvL vs. ETRL) as a fixed factor. A total of 128 participants took part in the study. The participants were mainly psychology students and randomly assigned to one of the two conditions. Table 1 provides an overview of the sample characteristics. To control for differences regarding reading

Table 1
Overview of the sample for the two conditions.

Sample characteristics	Total	Language Complexity		*t*-test	
		ConvL	ETRL	*t (df)*	*p*
Participants: *n*	128	64	64		
Women: *n* (%)	95 (74.2)	47 (73.4)	48 (75.0)		
Age: *M* (*SD*)	26.3 (10.6)	26.9 (11.3)	25.6 (10.0)	0.73 (126)	.47
Education: *M* (*SD*)	3.3 (0.6)	3.2 (0.7)	3.3 (0.6)	−0.9 (144)	.73
Literacy: *M* (*SD*)	16.2 (2.9)	15.9 (2.9)	16.5 (2.8)	−1.1 (126)	.70

Note. ConvL = conventional language; ETRL = easy-to-read language; n = number of participants; ♀(%) = percentage of women; Education: Likert scale from 1 to 4 (1 = no school-leaving qualification, 2 = primary education, 3 = secondary education, tertiary education); Literacy: C-test maximum points = 20.

abilities we conducted the C-test, which is an established measure to determine the literacy of participants (e.g., Grotjahn, 2010). This is a cloze test, in which people have to fill in missing words or letters to complete a text passage. The participants will receive a point for every correct answer and can achieve a maximum score of 20. There was no difference in literacy between the two experimental groups.

2.2. The web sites

In a first step, we built a web site of a fictitious leisure centre offering different activities, such as indoor pool, sauna, bowling and miniature golf. The text content was taken from different existing web sites of leisure centres and adapted accordingly. We used a fictitious name and location for the leisure centre to make the content anonymous. In a second step, a professional translator translated the whole text of the web site into ETRL following the guidelines for ETRL provided by the network for easy language (see http://www.leichtesprache.org/images/Regeln_Leichte_Sprache.pdf). Afterwards, we duplicated the web site and implemented the text in ETRL on the second web site. This resulted in two versions of the same web site, only differing in language complexity (i.e., content like pictures, general structure, and navigation was identical). Fig. 1 shows screenshots of both versions.

2.3. Manipulation of language complexity

A professional translator translated the text from ConvL into ETRL by applying the ETRL guidelines. For example, these rules suggest that foreign words and idiomatic expressions should be avoided, difficult words should be explained or expressed in a simpler way and only one statement per sentence should be made (see Appendix 1 for all rules). It is to mention that the ETRL version contained more text to convey the same information. This is because making only one statement per sentence results in total in more characters than when including several statements in one sentence. To verify the impact of our manipulation we conducted a pilot study, involving nondisabled users, users with cognitive impairments, and an automated readability test to determine language complexity. Afterwards, we did an additional manipulation check with the current sample. As part of the pilot study, we first showed the text of the leisure centre web site to 47 nondisabled users. Each participant was presented the version either in ConvL or in ETRL and had to rate the difficulty of the text on three items (i.e., Please rate the simplicity of the text; how well would a child in primary 3 understand the text; the sentences of the text are written in an easy language) using seven-point Likert scales. The results clearly revealed that nondisabled users perceived the text in ETRL ($M = 5.56$, $SD = 1.06$) to be easier than the one written in ConvL ($M = 4.08$, $SD = 1.06$) with a significant statistical difference, $t\ (45) = -5.317$, $p < .01$, $d = -1.58$. Second, since ETRL primarily aims to support people with problems in understanding written language we pre-tested the two web sites with five people who have had cognitive impairments. Each person had to read the web pages in both versions and had to decide whether he or she preferred the version in ConvL or in ETRL. All five participants preferred the version with ETRL and mentioned that the version with ConvL was too difficult to understand. Third, we conducted an automated readability test by using the German version of the Flesch-Reading-Ease readability formula (Amstad, 1978; Flesch, 1948). This formula allows an evaluation of the text difficulty using a score ranging from 0 to 100 (100 = very easy, 0 = very difficult). The comparison of the two text versions revealed a readability score of 52 (i.e., moderate difficulty) for the conventional version and a readability score of 67 (i.e., low difficulty) for the easy version. Finally, we conducted a manipulation check with the current sample by using the same items as in the pretest. The results indicated a significant difference between the ConvL ($M = 5.14$, $SD = 1.01$) and ETRL ($M = 5.84$, $SD = 0.98$) version in an expected direction $t\ (126) = -3.98$, $d = -0.71$. Overall, the results of the pilot study and of the manipulation check showed that the manipulation was successful.

2.4. Material

The participants used a Laptop (Dell Latitude E7240; 12˝) with an external mouse to explore the web site.

2.5. Search tasks and reading tasks

During the experiment participants had to solve three search tasks and three reading tasks (see Table 2 for an overview of the tasks). The search tasks comprised questions, which participants had to answer by detecting information on the web site (e.g., How much is the entrance fee). This required participants to search actively for information on the web site by clicking links, scrolling and screening text passages. The reading tasks required participants to read a particular part of the web site (e.g., Please read the text on the web page with the title “indoor pool”). For these tasks, participants did not need to navigate the web site. Performance measures for the two types of task differed slightly (see measures section 2.6).

2.6. Measures

2.6.1. Dependent variables

2.6.1.1. Performance measures. Seven performance variables were used, of which three can be considered efficiency and four effectiveness measures. The efficiency measures comprised (a) searching time (e.g. (Karreman et al., 2007) (b) reading time (e.g. Karreman et al., 2007), and (c) reading speed. The effectiveness measures consisted of (d) detection rate, (e) free recall of content (e.g. (Fletcher and Bloom, 1988; Lorch and Lorch, 1996; Nestojko et al., 2014; Whalon et al., 2015) (f) recognition of content (e.g., Chen et al., 2014; Karreman et al., 2007; Mangen et al., 2013), and (g) true/false statements (e.g., Ehrlich, 1999; Kasperski and Katzir, 2013; Meteyard et al., 2015). While some of these measures were taken when participants browsed the web site (i.e., a, b, c and d), the others (i.e., e, f, and g) were taken after browsing the web site. (a) Searching time measured the time (in seconds) needed to find a piece of information (see Table 2, task 1, 3, 5). (b) Reading time was a measure of the time (in seconds) spent reading a particular text passage (see Table 2, task 2, 4, 6). (c) Reading speed (milliseconds per character) was calculated by dividing the reading time by the total number of characters to read (see Table 2, task 2, 4, 6). (d) Detection rate referred to the percentage of successfully located information in search tasks (see Table 2, task 1, 3, 5). (e) Free recall of content, was measured with one task, in which participants had to name four activities described on the web site (see Appendix 2 for an overview

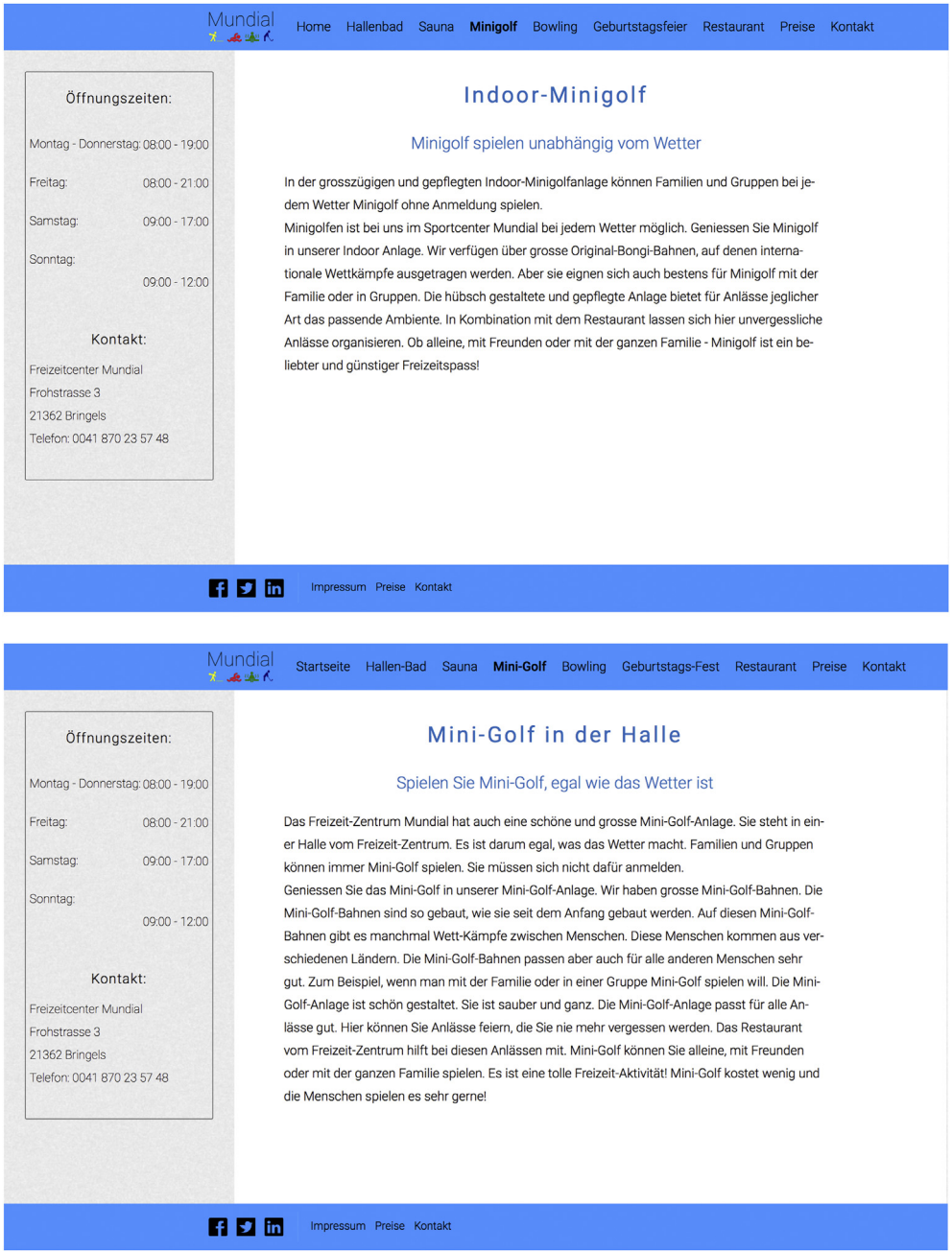

Fig. 1. Screenshots of a web page of the web sites used for testing (top: conventional language; bottom: easy-to-read language).

of the tasks e, f, & g). A maximum of four points was possible. (f) Recognition of content comprised two multiple-choice tasks, in which participants had to recognise content of the web site by choosing correct options from six answers (e.g., which of the following offers are provided in the swimming pool section?). A maximum of twelve points was possible. (f) True/false statements contained four statements (e.g., the water in the swimming pool has 30 °C). Participants had to give a true/false rating on a six-point Likert scale (i.e., "I am totally sure this

Table 2
Tasks to be completed on the web site.

Task	Type	Description
1	Search	How much is the entrance fee for the leisure centre if you want to stay for three hours?
2	Reading	Please read the whole text on the web page providing information about the indoor pool.
3	Search	The leisure centre offers a special birthday package for children. What is the minimum number of children that have to join the birthday party for this special offer to be valid?
4	Reading	Please read the whole text on the web page providing information about bowling.
5	Search	A restaurant in the leisure centre is called Sauna-Bar. Please put down the meals that this restaurant offers.
6	Reading	Please read the whole text on the web page providing information about the sauna.

Table 3
Dependent variables as a function of language complexity levels: means and standard deviations.

Measures	Conventional language Mean (SD)	Easy language Mean (SD)
Detection rate (%)	83.9 (16.7)	83.3 (17.0)
Detection time (s)	31.7 (14.3)	34.4 (17.3)
Reading time (s)**	42.1 (12.9)	59.4 (15.0)
Reading speed (ms/character)	53 (16)	56 (14)
Free recall (max. 4p)	3.7 (0.5)	3.7 (0.5)
Recognition (max. 12p)*	10.9 (1.0)	11.3 (0.8)
True/false statements (max. 20p)**	14.4 (2.8)	16.0 (3.5)
Perceived usability (1–5)	4.5 (0.3)	4.4 (0.4)
Perceived task load (1–20)	3.6 (1.7)	3.6 (2.3)
Perceived aesthetics (1–5)	5.3 (1.1)	5.0 (1.1)
Affective state: valence (1–9)	7.1 (1.4)	7.1 (1.4)
Affective state: arousal (1–9)	3.1 (1.7)	2.9. (1.6)
Text liking (1–7)**	4.8 (1.3)	3.5 (1.7)
Intention to revisit web site (1–7)**	6.1 (1.1)	5.3 (1.6)
Intention to visit leisure centre (1–7)	5.0 (1.6)	4.6 (1.7)
Information retention: free recall	3.4 (0.8)	3.5 (0.5)
Information retention: recognition	4.6 (1.0)	4.9 (1.0)
Information retention: true/false statements	11.3 (2.5)	12.1 (3.0)

statement is false" to "I am totally sure this statement is true"), with a maximum score of twenty points.

2.6.1.2. Subjective user reactions. A number of subjective user reactions were measured. (a) Perceived usability was assessed through the System Usability Scale (SUS; five-point Likert scale; Brooke, 1996). (b) To measure perceived taskload we used the NASA Task Load Index (NASA-TLX; 20-point Likert scale; Hart and Staveland, 1988). (c) Perceived aesthetics was measured by using the short version of the Visual Aesthetics of web sites Inventory (VisAWI-S; five-point Likert scale; Thielsch & Moshagen, 2011). (d) To measure affective reactions the Self-Assessment Manikin scale was used (SAM; ten-point Likert scale; Bradley and Lang, 1994). (e) To measure whether the participants liked the way in which the text on the web site was written, we used the following two items: I like the style in which text on the web site has been written; The writing style of the text on the web sites is appealing. Additionally, we asked (f) whether participants would consider revisiting the web site (I would revisit the web site if I had to find further information about the leisure centre) and (g) whether they would consider visiting the leisure centre (I would like to visit the leisure centre). These three items were rated on a seven-point Likert scale (totally disagree – totally agree).

2.6.1.3. Long-term information retention. Two weeks after the experiment, participants completed a test with six additional tasks comprising two free recall questions, one recognition task and three true/false statements. These tasks were structured similarly to the ones presented directly after testing (see Appendix 3).

2.7. Procedure

The experiment comprised the main testing session and a retention test administered two weeks after the main session. In the main session, participants had to solve six tasks on a laptop in the presence of the experimenter. A browser window was set up prior to the testing, which contained two tabs. One presented the instructions and the questionnaires while the other one showed the web site (either ConvL or ETRL) on which the participants had to complete the tasks. Participants started by reading the instructions on the screen and completed the literacy test. After they had completed the literacy test, they filled in the SAM to obtain a baseline of their current affective state. Afterwards they started with the tasks on the web site. When finishing the six tasks they had to complete the SAM again to assess possible changes in affective state. Thereafter, they filled in the other questionnaires to evaluate the web site (subjective usability, taskload, aesthetics etc.). Finally, they had to answer the free-recall, recognition of content and true/false tasks. Two weeks after the experimental session, the participants received an e-mail with a link to the retention test, which they could complete at home using their own computer.

2.8. Data analysis

A *t*-test for independent samples was used to compare the means of the different dependent variables for the two conditions (ConvL vs. ETRL). To measure changes in affective state, the baseline measure was used as a covariate in a one factorial analysis of covariance, with the post-test measurement of affective state being used as the dependent variable. This follows a procedure suggested by Dimitrov and Rumrill (2003).

Concerning the long-term retention test, we standardized the time lag between the testing session and the follow-up test by including only data from participants who responded on the 14th day after the testing session. This resulted in 29 participants being in the ConvL condition and 28 participants in the ETRL condition.

3. Results

3.1. Performance measures

3.1.1. Searching time

Table 3 shows the means and standard deviations of the dependent variables as a function of language complexity. Language complexity did not affect the detection time in search tasks, $t(126) = 0.97$, $p = .33$, $d = 0.18$.

3.1.2. Reading time

Participants who used the ETRL web site were slower in reading text than those who used the ConvL web site (see Table 3). This difference was significant, $t(126) = 6.96$, $p < .01$, $d = 1.25$.

3.1.3. Reading speed

The reading speed was similar for both web sites (see Table 3), with no significant difference, $t(126) = -1.35$, $p = .18$, $d = -0.24$.

3.1.4. Detection rate

Language complexity did not significantly affect the detection rate, $t(126) = -0.18$, $p = .86$, $d = -0.41$, (see Table 3).

3.1.5. Free recall of web site content

Participants in the ETRL condition showed similar scores in the free recall task as participants in the ConvL condition (see Table 3). The small difference between the two conditions was not significant $t(126) = -3.50$, $p = .72$, $d = -0.06$.

3.1.6. Recognition of web site content

The participants who used the ETRL web site showed higher scores in the recognition task than the participants in the ConvL condition (see Table 3). This difference was statistically significant $t(126) = -2.50$, $p = .01$, $d = -0.56$.

3.1.7. True/false statements

Using the ETRL web site led to higher scores in true/false statements than the ConvL condition (see Table 3). This difference was significant $t(126) = -2.990$, $p < .01$, $d = -0.53$.

3.2. Subjective user reactions

3.2.1. Perceived usability

Language complexity did not affect perceived usability, $t(126) = -1.76$, $p = .08$, $d = -0.32$.

3.2.2. Perceived taskload

There was no significant difference of taskload when comparing the two language complexity conditions, $t(126) = -0.20$, $p = .84$, $d = -0.04$.

3.2.3. Perceived aesthetics

Language complexity did not influence perceived aesthetics, $t(126) = -1.33$, $p = .19$, $d = -0.24$.

3.2.4. Affective state

Neither valence, $F(2, 125) = 0.04$, $p = .84$, $d = 0.00$, nor arousal, $F(2, 125) = 0.21$, $p = .89$, $d = 0.12$, was affected by language complexity.

3.2.5. Text liking

Participants who used the web site with ConvL gave higher ratings for text liking than participants who used the web site with ETRL (see Table 3). This difference in text liking was significant $t(126) = -5.00$, $p < .01$, $d = -0.90$.

3.2.6. Intention to revisit web site

The intention to revisit the web site was lower in the ETRL condition than in the ConvL condition (see Table 3). This effect was significant, $t(126) = 3.48$, $p < .01$, $d = 0.62$.

3.2.7. Intention to visit leisure centre

Language complexity did not significantly influence the intention to visit the leisure centre, $t(126) = 1.30$, $p = .20$, $d = 0.23$ (see Table 3).

3.2.8. Information retention

There was no significant effect on information retention. Neither in regard to free recall $t(94) = -1.25$, $p = .22$, $d = -0.23$, nor on recognition $t(94) = -1.61$, $p = .22$, $d = -0.33$, nor on true/false statements $t(94) = -1.57$, $p = .12$, $d = -0.32$ (see Table 3).

4. Discussion

This experiment examined the influence of the web accessibility criterion ETRL on nondisabled users. The results indicated that ETRL was associated with positive and negative consequences. On the one hand, ETRL increased recognition of web site content and scores in true/false statements. On the other hand, users in the ETRL-group needed more time to read the text and gave lower ratings for text liking and intention to revisit the web site than the ConvL-group. While the results are consistent with our assumptions regarding negative subjective reactions, the results concerning performance are not totally in line with our prediction. We assumed positive effects of ETRL on performance but the results showed a kind of a time–accuracy trade-off.

In line with previous work (Karreman et al., 2007), we showed that there are positive effects of ETRL on performance but detrimental effects on satisfaction ratings. Nondisabled users recognised more web site content and provided more correct answers to true/false statements but were less satisfied when using the ETRL version. This indicated that participants better processed information when interacting with simplified text. This is a highly relevant finding, which was also found in other areas such as designing medical information sheets for patients (e.g., Coyne et al., 2003; Root and Stableford, 1999; Van Weert et al., 2011). The drawbacks concerning lower text liking and lower intention to revisit the web site are of particular importance. For economic reasons, it is an important goal to increase user intention to revisit the web site (e.g., Vu and Proctor, 2011). Furthermore, the reduced text liking may have detrimental effects for the advertising effect of the web site. For the variables perceived usability, perceived aesthetics and affect, no effects of language complexity were found.

While ETRL improved some aspects of performance, there was also one aspect of performance that became worse under ETRL. Participants in ETRL needed more time to read text passages. The increase in reading time is not surprising because the manipulation of the language complexity resulted in more characters to read. This result is not in accordance with previous work, which showed no difference in reading time between ETRL and ConvL (Karreman et al., 2007). The reason for the different results may be due to differences in language characteristics and experimental manipulations. First (Karreman et al., 2007), manipulated text complexity and layout, but we only manipulated text complexity without changing the layout of the web site. Since in Karreman's study bullet points with shorter sentence structure were used to simplify the text, this may have led to a reduction in reading time. An important aspect of our manipulation was to make not more than one statement per sentence (see Appendix 1, recommendation 18) and to explain difficult words (see Appendix 1, recommendation 1). It is common to use longer sentences comprising a number of statements in the German language. Making only one statement per sentence results in total in a higher number of sentences than using sentences that include several statements. This is because for each statement a complete sentence structure is applied (e.g. at least one subject and one verb). Furthermore, the explanation of difficult words is often increasing the number of words used (see the example in Appendix 1, recommendation 1). Consequently, it is not surprising that the ETRL text was longer (ETRL: 2015 words, 295 sentences and 14095 characters; ConvL: 1279 words, 199 sentences and 9725 characters). Therefore, there seems to be a trade-off between simplified language and length of the text, which should be considered when applying ETRL. Although ETRL increased reading time, reading speed was not affected by the language manipulation. It is surprising that reading speed was the same for both groups although the ETRL text had a simpler structure. Previous research already showed that language complexity did not necessarily increase reading time (Rayner and Duffy, 1986). Rayner and Duffy (1986) showed that infrequent words and words with ambiguous meaning, as aspects of complex language, are crucial causes for increased reading time. Since the web site used in the present work described a leisure centre, the text content of the ConvL was not very complex but rather of medium complexity by using “everyday language” (e.g., describing restaurants or leisure activities). Therefore, the ConvL version comprised not many ambiguous and infrequent words, which might be an explanation of why reading speed was not affected by our language manipulation. In sum, some results of the present work correspond with previous work but others do not. Furthermore, the findings are partly in line with our assumptions. Although the findings of the present study provide important contributions to the field, there are some points to consider when interpreting the results.

4.1. Limitation, strengths and future research

Since the study focused on language complexity participants' reading abilities may have influenced the results. Although we aimed to recruit a heterogeneous sample, our participants were rather homogeneous, scoring rather high on the literacy test ($M = 16.2$; $SD = 2.9$; max. points = 20). The level of literacy may have influenced performance measures as well as the perception and evaluation of ETRL. We would expect that nondisabled people with lower literacy scores would have given higher ratings for text liking of ETRL because they might feel more comfortable when reading simpler text. Future studies should thus investigate nondisabled users with lower literacy scores.

Ratings of text liking may be affected by the kind of information presented. We used a web site with rather easy to understand content and a commercial purpose (i.e., advertising a leisure centre). The effects

of ETRL might have been different if we had used web sites presenting information of higher complexity (e.g., banks or insurances). It is conceivable that when the context was more complex and of higher importance (e.g., information about income), people would prefer ETRL. Future work should therefore also examine web sites with different complexity of information.

Since the ETRL-text was longer than the ConvL-text, it might be argued that the length of the text acted as a confounding factor. We would like to emphasise that, at least in German, an application of ETRL without changing the length of the text is very difficult, if not impossible, to achieve. Furthermore, in the present study we used an approach that is very similar to the procedure in practice (i.e., hiring a professional translator to translate the text). Nevertheless, future research should further investigate ETRL recommendations to extract the ones that have impact on the length of the text and try to find other solutions with less impact on the length of the text.

4.2. Implications for practitioners

This work provides important implications for practitioners. Possible side effects for nondisabled users should be considered when implementing accessibility criteria in web sites. The present results inform about potential benefits and drawbacks of ETRL for nondisabled users, which is highly relevant information for practitioners. Since ETRL seem to have beneficial effects on certain aspects of performance, the use of easy language may be especially interesting for web sites that aim to provide important information via text (e.g., medical, educational or governmental web sites). Using ETRL for such web sites may result in information being better processed by nondisabled users with similar benefits for users with disabilities. However, if a web site has mainly a commercial purpose (e.g., web sites that advertise products or services), one should use ETRL with some caution because nondisabled users may not like the writing style of the text and the advertising effect may be lost. Similar to other authors (i.e., Karreman et al., 2007), we suggest stating clearly the aim of the language style, which is to support people who have difficulties with reading. This might reduce negative user reactions. However, future research should empirically investigate whether such an intervention would be helpful. Another option would be to provide multiple versions of the same web site. It is already common to provide web sites in different languages (e.g., English, French and German). It would be possible to add ETRL as a further language option to choose from (e.g., Easy German), which would support users with various disabilities without interfering with the demands of nondisabled users.

Among the different issues in making web sites accessible, ERTL represents a bigger challenge than designing in accordance with the WCAG 2.0 (Schmutz et al., 2016). This is because previous studies on web accessibility showed only beneficial (and no detrimental) effects for nondisabled users when applying various disability-friendly design criteria, such as high contrast between text and background and visible focus (e.g., Schmutz et al., 2016). The accessibility criterion ETRL is more difficult to implement for nondisabled users since it involves a trade-off between benefits and potential drawbacks.

4.3. Conclusion

The present study showed that using ETRL on web sites entail benefits and drawbacks for nondisabled users. It seems to be important to consider contextual factors (e.g., purpose of web site, main user group, predominant type of tasks) when deciding on implementing ETRL as accessibility criterion. Referring to the present results and previous work on effects of accessibility on nondisabled users (e.g., Schmutz et al., 2016; Schmutz et al., 2017a; Schmutz et al., 2017b), it is to be concluded that considering accessibility in web site design may provide a wide range of benefits for nondisabled users. This is a very important finding, which needs to be disseminated to practitioners. Further work on effects of different aspects of accessibility on various user groups would increase the awareness of accessibility among practitioners, which would contribute to more equality and inclusion for web users.

Acknowledgement

We are grateful to Eva Weinzierl, Garance Lehmann, Lauriane Beffa, Natalie Pfister, Tatiana Koch, Veronica Gatti for their help in carrying out the research.

Appendix 1

Rules for ETRL provided by the network for ETRL

Use simple words. Explain difficult words.	
Bad example:	authorize
Good example:	allow
Bad example:	burdensome
Good example:	causing difficulty or worry
2. Use words that describe things exactly.	
Bad example:	public transport
Good example:	bus and train
3. Use common words.	
Bad example:	habitation
Good example:	house
4. Use always the same words for the same things.	
Bad example:	In a text about medicine, there is sometimes the word “tablet” used and sometimes the word “pill”.
Good example:	In a text about medicine, there is always the word “tablet” used.
5. Use short words.	
Bad example:	omnibus
Good example:	bus
6. Do not use abbreviations.	
Bad example:	FYI
Good example:	For your information
7. Use verbs.	
Bad example:	Tomorrow is the election of the president.

Good example:	Tomorrow, we will elect the president.
8. Use active voice.	
Bad example:	The paper is read by the woman.
Good example:	The woman reads the paper.
9. Do not use the genitive.	
Bad example:	The teacher's house.
Good example:	The house of the teacher.
10. Do not use the subjunctive.	
Bad example:	Maybe, it will rain tomorrow.
Good example:	It could rain tomorrow.
11. Do not use negations.	
Bad example:	Peter is not ill.
Good example:	Peter is healthy.
12. Do not use idiomatic expressions.	
Bad example:	A piece of cake.
Good example:	A task that can be accomplished very easily.
13. Do not use roman numerals.	
Bad example:	IV
Good example:	4
14. Do not use old dates.	
Bad example:	1877
Good example:	A long time ago.
15. Do not use large numbers.	
Bad example:	14.735 People
Good example:	A large number of people
16. Write telephone numbers with spaces between the numbers.	
Bad example:	00231 112211
Good example:	0 02 31 11 22 11
17. Do not use special characters.	
Bad example:	&
Good example:	and
18. Use short sentences. Make only one statement per sentence.	
Bad example:	There is a car, which is blue and black, beside the house.
Good example:	There is a car. The car is blue and black. The car is next to the house.
19. Use simple sentence structure.	
Bad example:	Together we drive to Spain.
Good example:	We drive to Spain together.
20. Address the readers directly.	
Bad example:	Tomorrow there will be an election.
Good example:	You can vote in the election tomorrow.
21. Do no use questions within the text.	
Bad example:	What might be the reason for this result? There are three possible options..
Good example:	There are three possible reasons leading to this result..
22. Do not use cross-references.	
Bad example:	see chapter 3 for a detailed description of “financial restrictions in international markets”
Good example:	“Financial restrictions in international markets” means …

Note. We adapted the recommendations and the examples from the German version. Since English and German differ considerably in grammar and sentence structure, some of the examples might not exactly apply to English.

Appendix 2

(d) Free recall of web site content (max. 4 points):

1. Please list 4 leisure activities that can be carried out in the leisure centre.
   Four activities should be listed, no options to choose from.

(e) Recognition of web site content (max. 12 points):

1. Which of the following is provided in the sauna area of the leisure centre?
   Six options to choose from: Finnish sauna, Russian sauna, stone sauna, steam shower, infrared sauna, and ecological sauna.
2. Which of the activities listed below are available in the swimming pool area of the leisure centre Mundial?
   Six options to choose from: Massage jets, jet stream pool, water slide, spring board, area for children, pool mattress.

(f) True/false statements (max. 20 points):

1. The water in the swimming pool has 30 °C.
   6-point Likert scale from "I am totally sure this statement is false" to "I am totally sure this statement is true".
2. On the web site it is recommended that people with heart problems should use the steam shower.
   6-point Likert scale from "I am totally sure this statement is false" to "I am totally sure this statement is true".
3. In the swimming pool there is an island of rocks.
   6-point Likert scale from "I am totally sure this statement is false" to "I am totally sure this statement is true".
4. Drinks are also offered in the bowling centre.
   6-point Likert scale from "I am totally sure this statement is false" to "I am totally sure this statement is true".

Appendix 3

Free recall of web site content (retention task, two weeks after testing, max. 5p):

1. Please write down the name of the leisure centre.
2. Please list 4 leisure activities that can be carried out in the leisure centre.
   Four activities should be listed, no options to choose from.

Recognition of web site content (retention task, two weeks after testing, max. 6p):

1. Which of the activities listed below are available in the swimming pool area of the leisure centre?
   Six options to choose from: Massage jets, jet stream pool, water slide, spring board, area for children, pool mattress.

True/false statements (retention task, two weeks after testing, max. 15p)

1. There is a Jacuzzi in the spa area.
   6-point Likert scale from "I am totally sure this statement is false" to "I am totally sure this statement is true".
2. On the web site it is recommended that people with heart problems should use the steam shower.
   6-point Likert scale from "I am totally sure this statement is false" to "I am totally sure this statement is true".
3. Drinks are also offered in the bowling centre.
   6-point Likert scale from "I am totally sure this statement is false" to "I am totally sure this statement is true".

References

Amstad, T., 1978. Wie Verständlich Sind Unsere Zeitungen? [How Readable Are Our Newspapers?] (Unpublished doctoral dissertation). University of Zürich, Zürich Switzerland.

Boldyreff, C., Burd, E., Donkin, J., Marshall, S., 2001. The case for the use of plain English to increase Web accessibility. In: Web Site Evolution, 2001. Proceedings. 3rd International Workshop on IEEE (pp. 42-48), https://doi.org/10.1109/WSE.2001.988784.

Bradley, M.M., Lang, P.J., 1994. Measuring emotion: the self-assessment manikin and the semantic differential. J. Behav. Ther. Exp. Psychiatr. 25 (1), 49–59. https://doi.org/10.1016/0005-7916(94)90063-9.

Brooke, J., 1996. SUS-A quick and dirty usability scale. Usability Evaluation in Industry 189 (194), 4–7.

Caldwell, B., Cooper, M., Reid, L.G., Vanderheiden, G., 2008. Web content accessibility guidelines 2.0, world wide web consortium (W3C) recommendation. Retrieved February 10, 2015, from http://www.w3.org/TR/2008/REC-WCAG20-20081211/ .

Chen, G., Cheng, W., Chang, T.-W., Zheng, X., Huang, R., 2014. A comparison of reading comprehension across paper, computer screens, and tablets: does tablet familiarity matter? Journal of Computers in Education 1 (2–3), 213–225. https://doi.org/10.1007/s40692-014-0012-z.

Coyne, C.A., Xu, R., Raich, P., Plomer, K., Dignan, M., Wenzel, L.B., et al., 2003. Randomized, controlled trial of an easy-to-read informed consent statement for clinical trial participation: a study of the eastern cooperative oncology group. J. Clin. Oncol. 21 (5), 836–842. https://doi.org/10.1200/JCO.2003.07.022.

Dimitrov, D.M., Rumrill Jr., P.D., 2003. Pretest-posttest designs and measurement of change. Work 20 (2), 159–165.

Ehrlich, M.F., 1999. Metacognitive monitoring of text cohesion in children. In: van Oostendorp, H., Goldman, S.R. (Eds.), The Construction of Mental Representations during reading (pp. 281–301). Lawrence Erlbaum Associates, Mahwah, NJ.

Ellcessor, E., 2014. Web accessibility myths as negotiated industrial lore. Crit. Stud. Mass Commun. 31 (2), 1–16. https://doi.org/10.1080/15295036.2014.91 9660.

Fajardo, I., Ávila, V., Ferrer, A., Tavares, G., Gómez, M., Hernández, A., 2014. Easy to read texts for students with intellectual disability: linguistic factors affecting comprehension. J. Appl. Res. Intellect. Disabil. 27 (3), 212–225. https://doi.org/10.1111/jar.12065.

Fajardo, I., Tavares, G., Ávila, V., Ferrer, A., 2013. Towards text simplification for poor readers with intellectual disability: when do connectives enhance text cohesion? Res. Dev. Disabil. 34 (4), 1267–1279. https://doi.org/10.1016/j.ridd.2013.01.006.

Farrelly, G., 2011. Practitioner Barriers to diffusion and implementation of web accessibility. Technol. Disabil. 23 (4), 223–232. https://doi.org/10.3233/TAD-2011-0329.

Flesch, R., 1948. A new readability yardstick. J. Appl. Psychol. 32 (3), 221–233. https://doi.org/10.1037/h0057532.

Fletcher, C.R., Bloom, C.P., 1988. Causal reasoning in the comprehension of simple narrative texts. J. Mem. Lang. 27 (3), 235–244. https://doi.org/10.1016/0749-596X(88)90052-6.

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? J. Bus. Ethics 113 (2), 363–375. https://doi.org/10.1007/s10551-012-1309-3.

Grotjahn, R., 2010. C-test: Contributions from Current Research, vol. 18 Peter Lang.

Hart, S.G., Staveland, L.E., 1988. Development of NASA-TLX (task load Index): results of empirical and theoretical research. Adv. Psychol. 52, 139–183. https://doi.org/10.1016/S0166-4115(08)62386-9.

Henry, S.L., 2006. Understanding web accessibility. In: Thatcher, J., Burks, R.M., Heilmann, C., Henry, S.L., Krikpatrick, A., Lauke, P.H. (Eds.), Web Accessibility: Web Standards and Regulatory Compliance (pp. 1-51). Apress. https://doi.org/10.1007/978-1-4302-0188-5_1.

Jacko, J.A., 2012. Human Computer Interaction Handbook: Fundamentals, Evolving Technologies, and Emerging Applications, third ed. CRC Press.

Karreman, J., Van Der Geest, T., Buursink, E., 2007. Accessible website content guidelines for users with intellectual disabilities. J. Appl. Res. Intellect. Disabil. 20 (6), 510–518. https://doi.org/10.1111/j.1468-3148.2006.00353.x.

Kasperski, R., Katzir, T., 2013. Are confidence ratings test-or trait-driven? Individual differences among high, average, and low comprehenders in fourth grade. Read. Psychol. 34 (1), 59–84. https://doi.org/10.1080/02702711.2011.580042.

Lorch Jr., R.F., Lorch, E.P., 1996. Effects of organizational signals on free recall of expository text. J. Educ. Psychol. 88 (1), 38. https://doi.org/10.1037/0022-0663.88.1.38.

Mangen, A., Walgermo, B.R., Brønnick, K., 2013. Reading linear texts on paper versus computer screen: effects on reading comprehension. Int. J. Educ. Res. 58, 61–68. https://doi.org/10.1016/j.ijer.2012.12.002.

Meteyard, L., Bruce, C., Edmundson, A., Oakhill, J., 2015. Profiling text comprehension impairments in aphasia. Aphasiology 29 (1), 1–28. https://doi.org/10.1080/02687038.2014.955388.

Nestojko, J.F., Bui, D.C., Kornell, N., Bjork, E.L., 2014. Expecting to teach enhances learning and organization of knowledge in free recall of text passages. Mem. Cognit. 42 (7), 1038–1048. https://doi.org/10.3758/s13421-014-0416-z.

Nurmela, K., Pirhonen, A., Salminen, A., 2013. Accessibility of public web services: a distant dream? In: Kotzé, P., Marsden, G., Lindgaard, G., Wesson, J., Winckler, M. (Eds.), Human-computer Interaction–INTERACT 2013 (pp. 566-578), https://doi.

org/10.1007/978-3-642-40483-2_40.
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. https://doi.org/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), https://doi.org/10.1145/990657.990660.
Rayner, K., Duffy, S.A., 1986. Lexical complexity and fixation times in reading: effects of word frequency, verb complexity, and lexical ambiguity. Mem. Cognit. 14 (3), 191–201. https://doi.org/10.3758/BF03197692.
Root, J., Stableford, S., 1999. Easy-to-Read consumer communications: a missing link in medicaid managed care. J. Health Polit. Pol. Law 24 (1), 1–26. https://doi.org/10.1215/03616878-24-1-1.
Ruth-Janneck, D., 2011. Experienced barriers in web applications and their comparison to the WCAG guidelines. In: Holzinger, A., Simonic, K. (Eds.), Information Quality in e-Health (pp. 283-300), https://doi.org/10.1007/978-3-642-25364-5_21.
Saggion, H., Štajner, S., Bott, S., Mille, S., Rello, L., Drndarevic, B., 2015. Making it simplext: implementation and evaluation of a text simplification System for Spanish. ACM Transactions on Accessible Computing (TACCESS) 6 (4), 14–36. https://doi.org/10.1145/2738046.
Schmutz, S., Sonderegger, A., Sauer, J., 2016. Implementing recommendations from web accessibility guidelines: would they also provide benefits to nondisabled users. Hum. Factors: The Journal of the Human Factors and Ergonomics Society 58 (4), 611–629. https://doi.org/10.1177/0018720816640962.
Schmutz, S., Sonderegger, A., Sauer, J., 2017a. Implementing recommendations from web accessibility guidelines: a comparative study of nondisabled users and users with visual impairments. Hum. Factors: The Journal of the Human Factors and Ergonomics Society 1–17. https://doi.org/10.1177/0018720817708397.
Schmutz, S., Sonderegger, A., Sauer, J., 2017b. Effects of accessible website design on nondisabled users: age and device as moderating factors. Ergonomics 1–13.
Thatcher, J., Burks, M.R., Heilmann, C., Henry, S.L., Kirkpatrick, A., Lauke, P.H., et al., 2006. Web Accessibility: Web Standards and Regulatory Compliance. Apress, Berkeley, CA.
Thielsch, M.T., Moshagen, M., 2011. Erfassung visueller Ästhetik mit dem VisAWI [Measuring visual aesthetics with the VisAWI]. Usability Professionals 9, 260–265.
Van Weert, J.C., Van Noort, G., Bol, N., Van Dijk, L., Tates, K., Jansen, J., 2011. Tailored information for cancer patients on the Internet: effects of visual cues and language complexity on information recall and satisfaction. Patient Educ. Counsel. 84 (3), 368–378. https://doi.org/10.1016/j.pec.2011.04.006.
Vu, K.-P.L., Proctor, R.W., 2011. Handbook of Human Factors in Web Design. CRC Press, Boca Raton, FL.
Whalon, K., Martinez, J.R., Shannon, D., Butcher, C., Hanline, M.F., 2015. The impact of reading to engage children with autism in language and learning (RECALL). Top. Early Child. Spec. Educ. 35 (2), 102–115. https://doi.org/10.1177/0271121414565515.
World Health Organisation, 2011. World Report on Disability. WHO Press, Malta.

Sven Schmutz is a PhD student at the University of Fribourg, Switzerland. He received his MSc 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.