Text Simplification and User Experience

Soussan Djamasbi[1]([✉]), John Rochford[2], Abigail DaBoll-Lavoie[1],
Tyler Greff[1], Jennifer Lally[1], and Kayla McAvoy[1]

[1] User Experience and Decision Making Research Laboratory,
Worcester Polytechnic Institute, Worcester, USA
{djamasbi,amdabolllavoie,tgreff,jmlally2,
kmmcavoy}@wpi.edu
[2] Eunice Kennedy Shriver Center, University of Massachusetts Medical School,
Worcester, USA
john.rochford@umassmed.edu

Abstract. Research provides ample evidence of the impact web page design
has on comprehension; and that Generation Y users are impatient and dislike
reading text. Yet there has been little research that focuses on content, in par-
ticular to examine the impact of text simplification on younger users' processing
of textual information. To address this need, we report the initial steps of a larger
research effort that focuses on developing a set of guidelines for designing
simple and effective text passages. Specifically, we compiled a set of existing
plain language rules and tested its effectiveness of conveying information to
Generation Y users. The results suggest the compiled set of rules can serve as an
appropriate tool for designing textual passages to reduce cognitive effort and
improve readability of textual content for Generation Y users. Also, the results
show that eye tracking serves as an excellent objective measurement for
examining the effectiveness of text simplification.

Keywords: Plain language standards (PLS) · Text simplification · Text
comprehension · Eye-Tracking · Cognitive effort · Performance · Generation Y

1 Introduction

According to a recent PEW report, 87 % of adults (18 +) and 95 % of teens (12 to 17)
in the United States (U.S.) are Internet users [9, 10]. An overwhelming majority of U.S.
users (90 %) report that Internet technologies have served as useful tools in their
personal lives. Another notable majority (76 %) think the Internet is beneficial for
society [6]. Because the Internet has become an essential source of information for the
majority of people in the U.S., effective communication of that information is of great
importance.

Many organizations pay close attention to how effectively they communicate
information to their users through their websites [3, 4]. A significant recommendation
is that web pages communicate information to users easily and quickly [7]. To address
this issue, a great number of investigators have examined the impact of visual
arrangements on the successful communication of information on webpages [5].
Despite that important information is often conveyed in textual format, eye-tracking

studies show viewers often pay little attention to text [4, 5]. For example, users read only about 20 % of text that is provided on a webpage [8]. These findings may suggest that simplifying text on websites is likely to improve the effectiveness of communication. However, little work has been done to provide a set of guidelines that (1) can significantly increase comprehension; and (2) can be tested for effectiveness with user subjective and objective reactions. We address this gap by testing the effectiveness of a comprehensive set of rules we compiled from various sources of plain language standards. We used eye tracking to capture user fixation during reading. Because fixation serves as a reliable measure of cognitive effort [4], eye tracking is a particularly useful tool to capture effort objectively and unobtrusively. We focus on Generation Y users only because research suggests this group of users particularly dislike reading text [4]. Hence, this generation is likely to benefit from this research.

1.1 Plain Language Guidelines

Plain language refers to clear writing that can be understood the first time it is read. In an effort to improve the public's understanding of the work U.S. federal agencies perform on the public's behalf, all were mandated to use plain language as of 2010 [11]. Such simplification is intended to improve access for the public, and is likely to improve engagement. For example, comprehensibility tends to be a universal characteristic of popular blogs, books, and novels [12]. Thus, a complete set of rules to compose simplified text could serve as an important tool for designing successful websites. To address this, we conducted a systematic search for plain language guidelines from various published sources. We consolidated the guidelines, including removing duplicates, to develop a comprehensive set for designing simple text (Table 1).

1.2 Research Motivation

While text simplification can improve accessibility of information for people with limitations in cognition and/or literacy, we argue that in today's digital world, text simplification will benefit all, not just those with such limitations. Simple text enables users to easily and quickly gather information. This is likely to be particularly important for Generation Y users, who tend to avoid reading textual information, find it boring to read long blocks of text, and prefer image based communication [4, 5]. In order to test the effects of simplified text on Generation Y users, we targeted college students and conducted two studies, which are described in the following section.

2   Method

We used plain language rules, listed in Table 1 below, to simplify text passages for two studies. In each study, we used two text passages (passage A and passage B) from two actual websites. Each text passage had two versions, original and simplified; hence each study had 4 different text passages (Original-A, Simplified-A, Original-B,

Table 1. Plain Language Guidelines

Rules	
• Identify and write for your audience	• Write short sentences (20-25 words), be succinct
• Avoid slang, jargon, colloquialisms, non-literal text	• Short paragraphs (no more than 150 words in 3-8 sentences)
• Use short, simple words (no more than ∼ 3 syllables)	• Use transition words in paragraphs (pointing words, echo links, explicit connectives)
• Use concrete, familiar words/combinations of words	• Check/use correct grammar and spelling
• Use "must" instead of "shall" ("must not" vs. "shall not")	• Use "you" and other pronouns to speak to the reader
• Use an active voice, simple present tense	• Organize document chronologically
• Avoid weak verbs (def: a verb that is	• Use lists
made past tense by adding -ed, -d, -t)	• Use tables to make complex material easier to understand
• Use parallel sentence structure	• Do not use ALL CAPS for emphasis
• Use positive terms (avoid "don't" or "didn't")	• Do not use underlining for emphasis
• Avoid multiple negatives ("don't forget to not…")	• Use bold and italics for emphasis
• Explain all acronyms/abbreviations and avoid if possible	
Sources	
• WebAIM, http://webaim.org/techniques/writing/	
• WebAIM – WAVE, http://wave.webaim.org/cognitive	
• Plain Language Association International, http://plainlanguagenetwork.org/plain-language/what-is-plain-language/	
• Plain Language Action and Information Network, http://www.plainlanguage.gov/site/about.cfm	
• U.S. Federal Plain Language Guidelines, http://www.plainlanguage.gov/howto/guidelines/FederalPLGuidelines/index.cfm	

Simplified-B). Two research team members first independently constructed the simplified passages. They then compared the two passages, and selected the best version, or constructed a best one from the two versions, for use as the final simplified text passage in the study.

2.1   Study 1

The objective of this initial study was to gather information about subjective user reactions to an initial set of simplified text. User reactions were captured via a short survey and interview questions. We used the plain language rules to simplify two text

passages (A or B) from two websites in the entertainment, movies, and games industries. These passages are displayed in Fig. 1. We asked 58 students from various disciplines in a university in the Northeast to read the two versions of each text passage, which were presented to them in random order displayed on web pages. We asked participants, after reading each passage, to evaluate (5-point rating scale: 1 = very hard, 5 = very easy) the readability of the text passage they just read. As displayed in Fig. 2 below, the simplified versions of the two text passages were perceived as very easy or "sort of" (somewhat) easy to read by the majority of the participants (91 % and 82 % for the movie and the game text passages respectively). The percentages of perceived
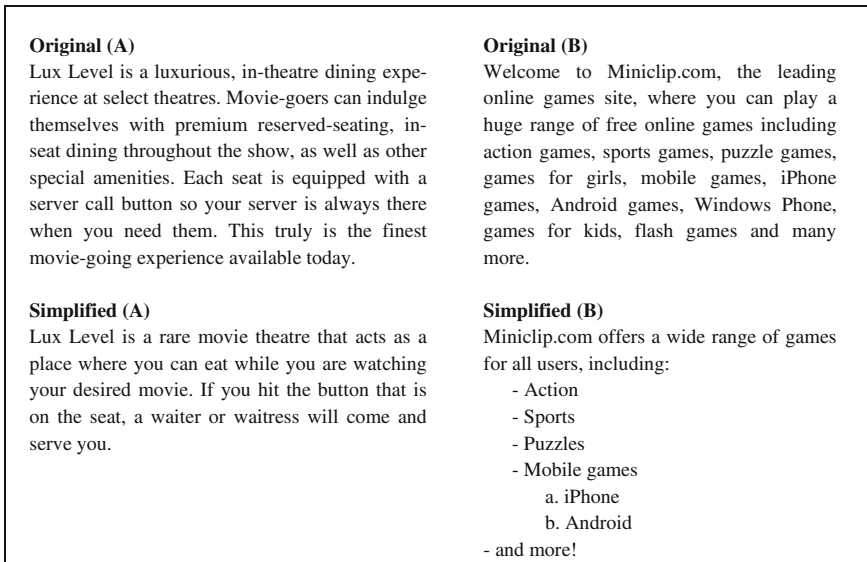
Original (A)
Lux Level is a luxurious, in-theatre dining experience at select theatres. Movie-goers can indulge themselves with premium reserved-seating, in-seat dining throughout the show, as well as other special amenities. Each seat is equipped with a server call button so your server is always there when you need them. This truly is the finest movie-going experience available today.

Simplified (A)
Lux Level is a rare movie theatre that acts as a place where you can eat while you are watching your desired movie. If you hit the button that is on the seat, a waiter or waitress will come and serve you.

Original (B)
Welcome to Miniclip.com, the leading online games site, where you can play a huge range of free online games including action games, sports games, puzzle games, games for girls, mobile games, iPhone games, Android games, Windows Phone, games for kids, flash games and many more.

Simplified (B)
Miniclip.com offers a wide range of games for all users, including:
  - Action
  - Sports
  - Puzzles
  - Mobile games
      a. iPhone
      b. Android
- and more!

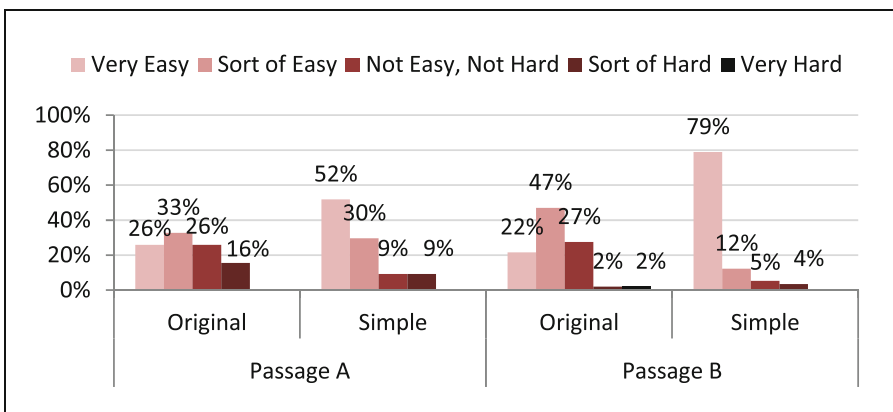**Fig. 1.** Text passages used in Study 1

Fig. 2. Survey Results

very easy or somewhat easy ratings were much lower for the original versions of the two text passages (69 % and 59 % respectively).

The results of a t-tests, as displayed in Table 2, showed that text, simplified with plain language rules for both versions, were reported significantly (p = 0.004 and p = 0.000) easier to read than the original text.

Table 2. t-tests Comparing Text Passages

	Original text	Simplified text
Passage A	3.69 (1.03)	4.24 (0.97)
Passage B	3.84 (0.86)	4.67 (0.74)
	df = 110, t Stat = 2.91, p = 0.004	*df = 106, t Stat = 5.36, p = 0.000*

Participants' comments, collected in a brief interview after the participants completed rating the text passages, supported the above survey analysis. They indicated participants preferred reading simplified text. The comments describing reactions to the simplified text included statements such as "details are there and understandable," "extremely basic description," "quick and to the point," "shorter sentences," "not wordy," "no fluff," and "can easily skim and still understand the information". The comments describing reactions to the original text included statements such as "takes more time to read and understand," "easy to lose track," "zoned out while reading it," "very dense, many will glance over it," "not quick or to the point," and "used unnecessary words".

These results suggest the rules in Table 1 may serve as an appropriate tool for developing easy to read text passages for Generation Y users.

2.2   Study 2

To examine the impact of plain language rules (Table 1) on effective communication, we conducted a laboratory eye tracking study to objectively measure effort, comprehension, and performance. Effort plays an important role in how people use information systems [1]. Thus, the degree to which people are willing to expend effort when reading text is likely to affect comprehension of the available information.

We recruited 18 participants from the same pool as that of the first study. We focused on websites that provided information about health and wellness. We selected two health-related text passages from two blog posts (Fig. 3). These two text passages were longer than the text passages used in Study 1. One provides information about the importance of a healthy breakfast in maintaining healthy weight. The other provides tips for taking action against indoor and outdoor allergies. We created simplified versions of each using the same procedure described for Study 1. To prevent a possible comprehension-bias effect that could have occurred had participants read original text passages followed by simplified versions of them, we presented only one version (Original or Simplified) of each passage to each participant.

We showed four pages to each participant. Page 1 displayed a text passage. We asked participants to read it, then to navigate to a second page, by invoking a "next" button, once they were finished reading.

Page 2 displayed a set of questions related to the passage just read, and the passage itself. The display of both the questions and the related text passage enabled participants to refer easily to the passage while answering the questions. We asked participants to answer the questions, and then to navigate to a third page via a "next" button.

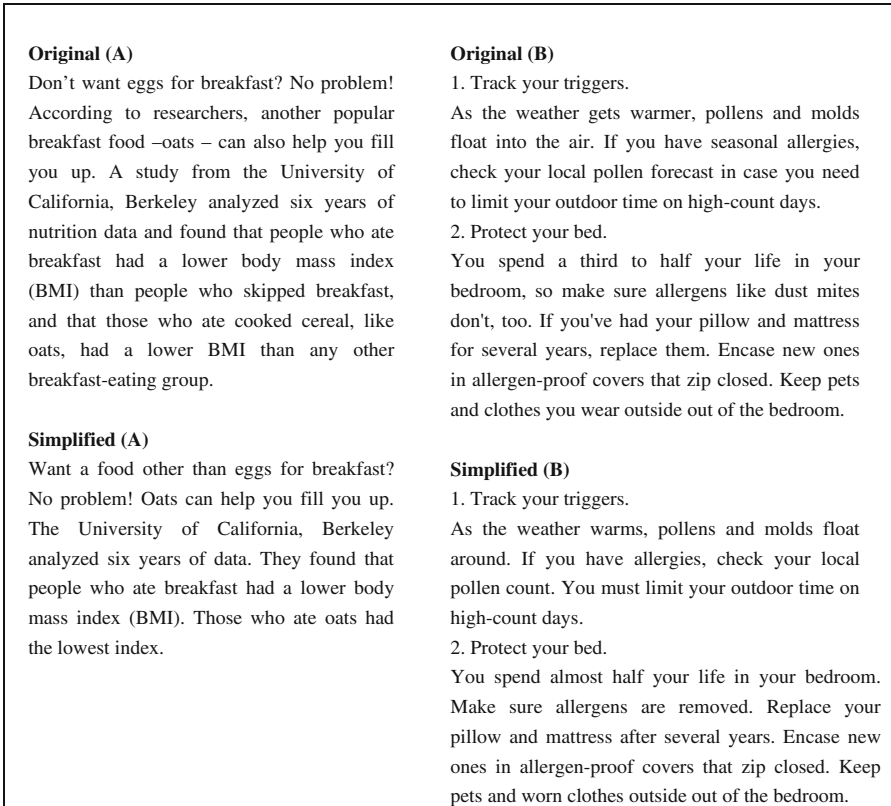Pages 3 and 4 operated as pages 1 and 2, but with a different text passage (Fig. 3.)

Original (A)

Don't want eggs for breakfast? No problem! According to researchers, another popular breakfast food –oats – can also help you fill you up. A study from the University of California, Berkeley analyzed six years of nutrition data and found that people who ate breakfast had a lower body mass index (BMI) than people who skipped breakfast, and that those who ate cooked cereal, like oats, had a lower BMI than any other breakfast-eating group.

Simplified (A)

Want a food other than eggs for breakfast? No problem! Oats can help you fill you up. The University of California, Berkeley analyzed six years of data. They found that people who ate breakfast had a lower body mass index (BMI). Those who ate oats had the lowest index.

Original (B)

1. Track your triggers.

As the weather gets warmer, pollens and molds float into the air. If you have seasonal allergies, check your local pollen forecast in case you need to limit your outdoor time on high-count days.

2. Protect your bed.

You spend a third to half your life in your bedroom, so make sure allergens like dust mites don't, too. If you've had your pillow and mattress for several years, replace them. Encase new ones in allergen-proof covers that zip closed. Keep pets and clothes you wear outside out of the bedroom.

Simplified (B)

1. Track your triggers.

As the weather warms, pollens and molds float around. If you have allergies, check your local pollen count. You must limit your outdoor time on high-count days.

2. Protect your bed.

You spend almost half your life in your bedroom. Make sure allergens are removed. Replace your pillow and mattress after several years. Encase new ones in allergen-proof covers that zip closed. Keep pets and worn clothes outside out of the bedroom.

Fig. 3. Text passages used in Study 2

2.3   Measures

Fixations, or relatively steady gazes that are at least 60 ms long, are reliable indicators of cognitive processing when reading text passages [3]. We thus used gaze analysis to measure cognitive effort. We used the Tobii × 300 eye tracker to capture participants' eye movements unobtrusively. Before starting the task, each participant went through a brief calibration procedure, which enabled us to collect the participant's eye movements. We used the accuracy of the answers to gauge performance.

3   Results

An overall look at the data showed more correct answers to questions related to simplified text passages. As shown in Fig. 4, the means of correct answers for both simplified text passages were larger than those for the original text passages (1.78 vs. 1.61). Additionally, 83 % of participants were able to answer all questions correctly for the simplified text, while a lower number (67 %) were able to do the same for the original text passages (Fig. 4b). We observed an upward trend in performance for both simplified text passages. However, this trend was more nuanced for passage A (Fig. 4).
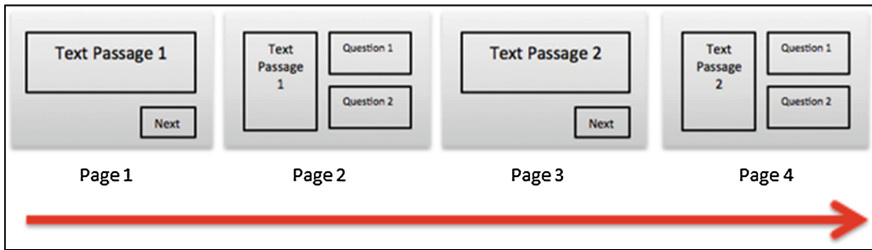
Fig. 4. The flow of pages presented to each user. Latin square design was used to avoid order effect.

Eye tracking data revealed that participants read the simplified text passages in a shorter time regardless of whether they were reading passage A or B (Fig. 5). This trend suggests participants processed simplified text passages more efficiently, with less effort. Unsurprisingly, the results show participants took longer to read passage B, which was longer than passage A.

Looking at the distribution (Fig. 6) on fixations of text passages and questions on the pages (2 & 4) that contained both of these components, we can see participants attended more to answering questions than to reading the text passages. This is not surprising because participants were able to read the text passages on the previous pages (1 & 2), and therefore were using the text passages more as a reference for answering questions.

The data show that the difference between the distribution for simplified and original versions was more pronounced on the set of passage As. Fixation duration was largest when participants responded to the questions for the simplified version of passage A, and shortest when they referred to its related text passage on the same screen. These results, along with performance data (100 % accuracy for simplified passage A), suggest that the simplified version of passage A was processed more efficiently and effectively than the other passages. While the reported differences in this section were not significant, together, the results indicated an upward trend in efficiency and performance for the simplified text (Fig. 7).

a. Average number of correct answer (maximum 2 per passage)

b. Percentage of correct answers for a passage
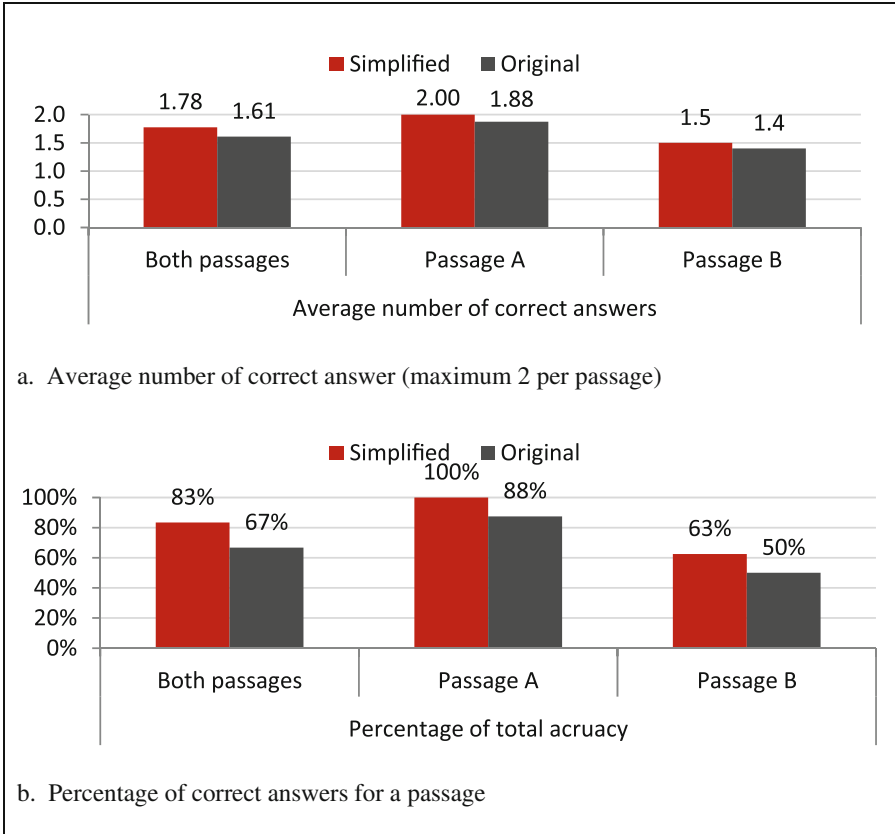
Fig. 5. Performance
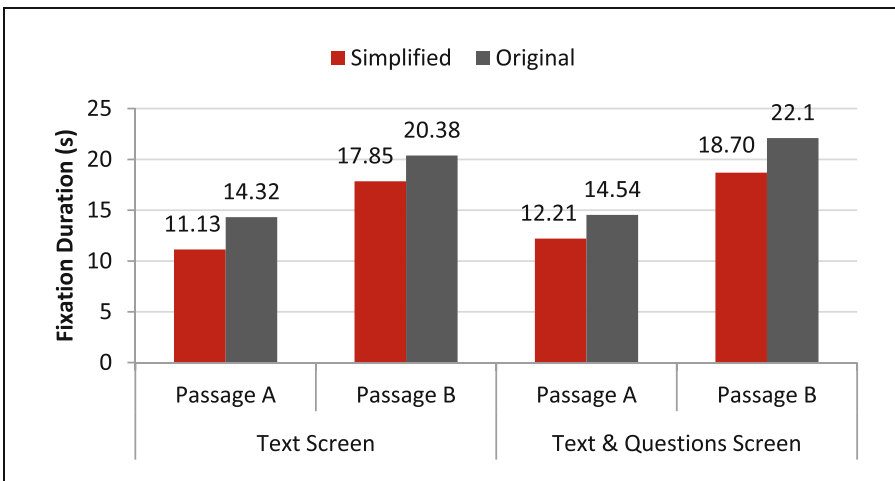
Fig. 6. Average of participants' fixation duration on each of the two screens
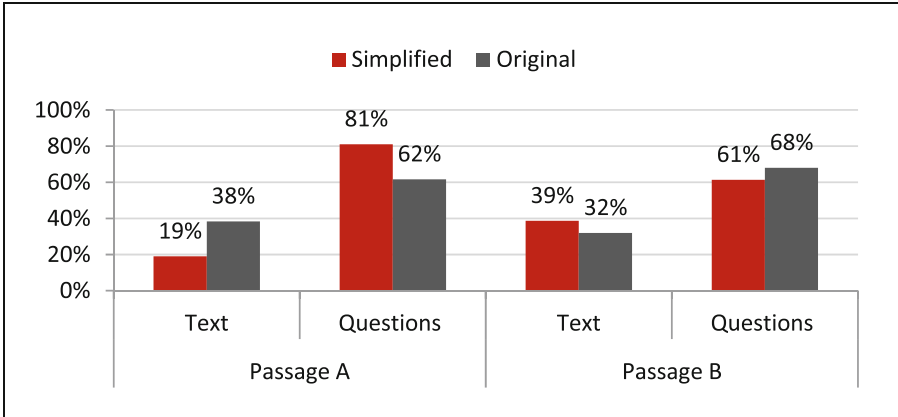
Fig. 7. Distribution of fixation duration on text & questions screen

4   Discussion

The objective of this study was to examine the impact of text simplification on effective communication. To do so, we compiled a comprehensive set of rules obtained from existing sources that define plain language standards. We used the compiled data set to simplify a set of text passages from actual websites.

We then conducted a preliminary study to see whether users responded well to simplified text. The results of this initial study showed that people provided significantly better comprehension ratings for the simplified text. Their comments indicated that they preferred reading simplified text on web pages.

In the second study, we tested objective reactions to another set of passages that were slightly longer in length. The eye movement data suggested that participants processed simple text passages more efficiently. The performance data showed that, on average, people answered more questions correctly when reading simplified text. Hence, the performance data suggested that simplified passages were easier to understand. Together, these results suggested that text simplification improves the effectiveness of information communicated to Generation Y users.

The results of this study also suggest that our compiled set of standards provided an appropriate initial set of guidelines for constructing simplified text that can improve reading comprehension and performance.

The results of this study have important theoretical and practical implications. Given the importance of cognitive effort in effective usage of computerized information tools [1], the results provide a theoretical direction for further development of text simplification guidelines. Because the results show that Generation Y users prefer simplified text, they also provide insight for HCI research that focuses on younger users. Generation Y's eye movement data generally reveal an "impatient" pattern of viewing [2, 4, 5]. This may be because younger users dislike reading text [4, 5]. Thus having guidelines for simplifying text is of great importance in designing effective communication for Generation Y users. From a practical point of view, the results

provide guidelines that can be used to design websites with effective communication in mind. Given today's crowded web environment, providing effective communication could serve as invaluable strategy to drive behavior and improve market share.

As with any study, the results of our study are limited to the task and to the study setting. The sample size is yet another limiting factor in our study. Future studies with larger sample sizes and different tasks are needed to verify our results and to extend their generalizability.

5   Conclusion

The objective of this study was to compile a set of standards for simplifying textual information, and to test its effectiveness on communicating textual information with Generation Y users. Testing the original passages and the simplified passages together allowed for a direct comparison regarding the effectiveness of the compiled set of standards. The results show that the compiled set of rules in our study has the potential to effectively reduce users' cognitive effort and thus improve their performance when reading text. The results also show that eye tracking serves as an excellent tool for this line of research because it can capture effort objectively, continuously, and unobtrusively. The results have important theoretical implications for HCI researchers who study the impact of cognitive effort on comprehension and performance. The results have also important implication for designers who focus on Generation Y users.

References

1. Djamasbi, S.: Does positive affect influence the effective usage of a decision support system? Decis. Support Syst. **43**(4), 1707–1717 (2007)
2. Djamasbi, S., Hall-Phillips, A.: Eye tracking in user experience design, pp. 27–43. Morgen Kaufmann (2014). http://digitalcommons.wpi.edu/uxdmrl-pubs/42
3. Djamasbi, S., McAuliffe, D., Gomez, W., Kardzhaliyski, G., Liu, W., Oglesby, F.: Designing for success: creating business value with mobile user experience (UX). In: Nah, F.F.-H. (ed.) HCIB 2014. LNCS, vol. 8527, pp. 299–306. Springer, Heidelberg (2014)
4. Djamasbi, S., Siegel, M., Skorinko, J., Tullis, T.: Online viewing and aesthetic preferences of generation Y and the baby boom generation: testing user web site experience through eye tracking. Int. J. Electron. Commer. **15**(4), 121–158 (2011)
5. Djamasbi, S., Siegel, M., Tullis, T.: Generation Y, web design, and eye tracking. Int. J. Hum Comput Stud. **68**(5), 307–323 (2010)
6. Fox, S., Rainie, L.: The Web at 25 in the US, Pew Research Center (2014). http://www.pewinternet.org/2014/02/27/the-web-at-25-in-the-u-s/
7. Krug, S.: Don't Make Me Think: A Common Sense Approach to Web Usability. Pearson Education, India (2005)
8. Nielsen, J.: How little do users read? Nielsen Norman Group (2008). http://www.nngroup.com/articles/how-little-do-users-read/
9. PEW: Teens Fact Sheet. Pew Research Center (2012). http://www.pewinternet.org/fact-sheets/teens-fact-sheet/

10. PEW: Internet Overtime. Pew Research Center (2014). http://www.pewinternet.org/datatrend/Internet-use/Internet-use-over-time/
11. The Plain Language Action and Information Network (PLAIN), Plain Language: Improving Communications from the Federal Government to the Public (2011). http://www.plainlanguage.gov/
12. Snow, S.: This Surprising Reading Level Analysis Will Change the Way You Write, The Content Strategist (2015). https://contently.com/strategist/2015/01/28/this-surprising-reading-level-analysis-will-change-the-way-you-write/