A preliminary examination of the effect of white and blue backgrounds on Web-based English listening tests

Atsuko K. Yamazaki\textsuperscript{a}\textsuperscript{*}, Hitomi Shimada\textsuperscript{a} and Kaoru Eto\textsuperscript{b}

\textsuperscript{a}College of Engineering, Shibaura Institute of Technology, 307 Fukasaku, Minuma-ku, Saitama-shi, Saitama, Japan
\textsuperscript{b}Faculty of Engineering, Nippon Institute of Technology, 4-1 Gakuendai, Miyashiro-cho, Minami Saitama-gun, Saitama, Japan

Abstract

Previous studies conducted by Yamazaki and Eto on the background colors of on-line tests indicated that a combination of black text and a white background color was not considered preferable for Web-based tests (WBTs). In this study, the authors conducted an experiment to see if a background color can affect the scores of a Web-based English listing test. Thirty three subjects in their twenties took Web-based English listening tests and non-linguistic tests with blue and white background colors with black text. The average test scores of the subjects showed no significant differences between the blue and white backgrounds. The brain functions of the subjects were also examined by observing relative changes in hemoglobin (Hb) concentrations in their brains with near-infrared spectroscopy. Two dimensional images of the Hb concentration changes obtained in the experiment indicated that areas in the brain associated with language processing tended to have higher Hb concentrations while the subjects were taking the listening test with both background colors. On the other hand, areas in the brain related to the frontal eye field were observed to be more active while they were taking the test with the white background than with the blue background, even though the questionnaire result obtained from the subjects showed that they felt more tired when they took the tests with the combination of black text and a blue background.

© 2013 The Authors. Published by Elsevier B.V. Open access under CC BY-NC-ND license.
Selection and peer-review under responsibility of KES International

Keywords: Background colors, Web-based test, English listening, test performance, brain functions, near-infrared spectroscopy, Brodmann Area.

1. Introduction

Web-based tests (WBTs) have been widely used because they provide cost efficiency as well as immediate feedback on test takers’ performance. In particular, web-based testing to assess English proficiency, such as the TOEFL® test and the TOEIC® test, has been utilized by many schools and companies. In spite of the

\textsuperscript{*} Corresponding author. Tel.: +0848-687-5724; fax: +08-48-687-5013.
\textit{E-mail address:} atsuko@sic.shibaura-it.ac.jp
increasing popularity and influence of WBTs, not enough attention has been paid to the visual designs of a WBT in relation to test takers’ performance. Some studies have been conducted to examine the effects of color combinations used for web-based teaching and learning materials [1][2]. However, not much research has been done to investigate whether differences in background colors used on Web-based listening tests affect the performance of test takers. Many English WBTs use black text against a white background, resembling to conventional paper-based tests. The text-background combination has never been evaluated in terms of test takers’ performance on a Web-based English listening test.

Many studies on colors used in web design have pointed out that the combination of background and text colors is important for performing tasks on Web pages in terms of readability. From their experimental results, Hall and Hanna pointed out that greater contrast ratios between background and font colors lead to better readability. Their results, however, showed that the combination of background and text colors did not significantly affect users’ retention [3]. On the other hand, other studies have found that a luminance contrast ratio between the text and its background values can affect readability [4][5][6]. According to a study of background colors conducted by Mehta and Zhu [7], red is beneficial for some kinds of mental processing while blue is better for others tasks. In their experiments, the results suggested that red backgrounds were better for tasks that require an attention to detail, and blue enhanced performance on approach-based and exploratory tasks. Their experiments showed that the subjects performed better on a word-recall task and a proof-reading task when the screen background was red, while the subjects came up with better quality and more creative ideas for things to do with a brick when the screen was blue, rather than red.

2. Previous studies and the purpose of this study

Yamazaki examined whether the background colors of Web-based tests have significant influence on the scores of test takers [8]. The results of the experiments in these studies demonstrate that the background color of a WBT can affect the performance of a test-taker. The results also suggest that blue colors may be better for the background color of a WBT when test sentences are displayed in black. Yamazaki’s study indicated that the background color of a WBT made a difference to the concentration level of a test taker, and that primary colors with low luminance ratios cause test takers to lose their concentration [8]. On the other hand, the results from the study suggested that the color preference of the user was not an important factor in choosing the best background color for a WBT [8], although Hall and Hanna demonstrated that preferred color would lead to higher ratings of behavior intention [2].

Yamazaki and Eto conducted an experiment to investigate if test takers’ scores differs, depending on the characteristics of white and blue background colors of Web-based tests [9]. They also examined how background colors can affect the functions of WBT takers’ brains by observing relative changes in hemoglobin (Hb) concentrations in their brains by using near-infrared spectroscopy. The average WBT scores of the linguistic and non-linguistic tests for the blue background were higher than those for the white background among the subjects. Yamazaki and Eto also found the same tendency in their preliminary study, in which the effect of eight background colors on the scores of a computer-based English grammar test was examined [10]. Table 1 summarizes the hexadecimal color codes of white and blue colors used in these studies, as well as the average percentage of questions answered correctly by the subjects for the blue and white backgrounds. The table also shows the numbers of English grammar questions that the subjects answered in the studies [8][9][10].

Since the blue colors resulted in the two highest test score averages in Yamazaki’s study [8] and the average test scores of the blue background were much higher than those of the white background in other previous studies by Yamazaki and Eto [9][10], a combination of black text and a white background color is not considered preferable for Web-based English grammar tests. In these studies, the brain regions associated with linguistic tasks was observed to be more active while the subject were taking the linguistic tests with the blue
background [9][10]. Therefore, we suspect that neurological factors associated with color characteristics can explain the differences among the test score averages for these background colors.

Table 1. Hexadecimal codes of blue and white backgrounds and the average percentage of questions answered correctly for the white and blue backgrounds in the previous studies [8][9][10]

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Blue</td>
<td>#0000FF</td>
<td>31.42 % (28)</td>
<td>59.17 % (24)</td>
<td>57.80 % (3)</td>
</tr>
<tr>
<td>White</td>
<td>#FFFFFF</td>
<td>30.24 % (31)</td>
<td>49.72 % (24)</td>
<td>34.27 % (7)</td>
</tr>
</tbody>
</table>

In this study, the authors investigated whether a background color can also affect the scores of Web-based English listening tests in a way similar to how they affected the results in the previous studies of English grammar tests. In order to see if the activation of linguistic regions in the brain differs in relation to the background color of the English listening test, we observed the brain functions of WBT takers who were answering English listening questions. We examined the activities of WBT test takers’ brains while they were taking a Web-based English test with blue and white screen background colors.

3. Experiment

The experiment was conducted with two sets of web-based tests, which consisted of English listening and non-linguistic test questions. The latter served as rest tasks. A white background was used for one test set and a blue background was chosen for the other test set. The text color of all questions was black for both sets. A total of 33 subjects participated in this experiment. In order to identify which part of WBT test taker’s brain is activated while the test taker is answering a listening question with each background color, we measured relative changes in blood hemoglobin (Hb) concentrations in the brain of the test taker by using a near-infrared spectroscopy (NIRS) system.

3.1. Web-based test

For the experiment, two sets of Web-based English listening tests with white and blue background colors and black text were developed. Both test sets had 15 English listening questions that were very similar to questions in Part II of the TOEIC® Test, which is a standardized English examination. The TOEIC® Test is a multiple-choice exam that consists of two parts with 100 questions each: the Listening Section and the Reading Section. The Listening Section is divided into four parts. The second part of the Listening Section is composed of 30 question-response questions. The Web-based English tests constructed for the experiment had 15 listening questions of the same type and of a similar difficulty level as Part II of the TOEIC® Test. The experimental tests were presented to the test takers and a response for each question was indicated by selecting one of the numbers, 1 through 3, for the answer choices with the pull-down menu. Examples of the test pages with the white background and the blue background are shown in Figure 1.

In both sets of the WBTs, a non-linguistic task was presented at the beginning of the test and after every three English grammar questions to serve as a rest task. All the rest task pages were designed in the same way.
Circles, stars and triangles in black color were randomly placed on single-color backgrounds. A test taker was directed to count the number of circles on the page and select a choice corresponding to the number of circles from the pull down menu. Examples of the rest task pages with the white background and the blue background are shown in Figure 2.

3.2. Method

The 33 subjects were university students and graduate school students in their twenties. Their first language was Japanese. The subjects consisted of 26 males (78.8%) and seven females (21.2%) and they were all right-handed. None of the subjects was reported to have a color vision deficiency at the time of the experiment. They had all taken a paper-based TOEIC® Bridge Test or TOEIC® Test before the experiment. The subjects were given instructions for taking the WBTs prior to undertaking the test. All the subjects took the tests with black characters and blue and white backgrounds. Test scores obtained from the subjects were analyzed to see if there were any differences between the two test sets with the different background colors. We used an NIRS system developed by Hitachi Medical Corporation (EGT 4000, 52 channels) to observe and record hemoglobin
concentrations in the brain of every test taker. Figure 3 shows the setup of the brain function measurement in this experiment. Figure 3 (a) is the photo of one of the subjects taking the WBT with the white background while wearing a headgear for applying NIRS probes to his forehead. Figure 3 (b) shows the probe arrangement used in the experiment.

The headgear included 52 optical source-detector channels to monitor relative changes in hemoglobin concentration in subject’s brain. The channels only covered the left hemisphere of the brain, including a part of the frontal region of prefrontal cortex, Broca’s Area and Wernicke's Area. The covered regions roughly corresponded to Brodmann Areas (BA) 8, 9, 10, 19, 22, from 39 to 46 and 52. Figure 3 (b) shows the positions of NIRS probes on subject’s head. In order to assess the activation of the brain functions associated with these areas, the blood hemoglobin concentrations of each subject were observed from the beginning of the tests until the subject completed the last listening question, and relative changes in oxy-hemoglobin, deoxy-hemoglobin, and total hemoglobin (oxy-Hb, deoxy-Hb, total Hb) concentrations from the 52 channel points were simultaneously measured and recorded for each subject.

After finishing the tests, each subject responded to a questionnaire designed to determine whether subjects’ concentration levels and feelings about the test had affected his test score. One of the questions asked the subjects if they felt tired from taking a test with each background color by using a five-point scale from 1=“did not feel tired at all” to 5=“felt very tired”. The subjects also indicated their concentration levels at the time of taking the test using a five-point scale from 1=“could not concentrate at all” to 5=“could concentrate very well”. They also responded to questions about how difficult they felt the test was and how tired they felt after the test with another five point scale.

**Fig. 3.** Setup of the brain function measurement: (a) a subject taking the Web-based listening English test with optical topography probes applied to his forehead; (b) the probe arrangement used in the experiment

### 4. Result

The average test scores obtained for each background color are summarized in Table 2. The average scores for both the English and the non-linguistic tests taken with the blue background were a little higher than those for the white background. Even though the difference was not as significant as the results from the previous experiments on the grammar tests [9][10], the result from this experiment on the listening tests coincided with the finding of the studies of Web-based grammar tests, wherein the blue backgrounds resulted in higher
average scores than those with the white backgrounds. The results for the concentration level and the degree of fatigue for the subjects were more positive for the white background than the blue one. The average scores of the concentration level and the degree of fatigue obtained from the subjects for each background color are summarized in Table 3.

Table 2. Average test scores of the English listening and circle counting tests for the blue and white backgrounds

<table>
<thead>
<tr>
<th>Background Color</th>
<th>#of Subjects</th>
<th>Average % of correct answers of English listening test</th>
<th>Average % of correct answers of circle counting test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blue</td>
<td>33</td>
<td>90.30%</td>
<td>72.32%</td>
</tr>
<tr>
<td>White</td>
<td>33</td>
<td>89.09%</td>
<td>68.89%</td>
</tr>
</tbody>
</table>

Table 3. Concentration and tiredness levels for the blue and white backgrounds indicated by the subjects

<table>
<thead>
<tr>
<th>Background Color</th>
<th>#of Subjects</th>
<th>“Could concentrate” very well &amp; well (%)</th>
<th>“Felt very tired ” &amp; “Felt tired” (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blue</td>
<td>33</td>
<td>21.88</td>
<td>71.88</td>
</tr>
<tr>
<td>White</td>
<td>33</td>
<td>50.00</td>
<td>28.13</td>
</tr>
</tbody>
</table>

The oxy-Hb concentration changes in subjects’ brains recorded during the experiment were mapped onto two-dimensional images of the brains. The images were also analyzed in relation to the functions of brain regions. Figure 4 shows a typical example of Hb concentration images created from the Hb measurements while the subjects were taking the Web-based listening tests with the blue and white backgrounds. The both images show relative changes in oxy-Hb concentrations in the brain of one of the subjects (Subject A), right after the subject started to listen to an English sentence. The TOEIC score of Subject A was reported as 500, which roughly corresponds to the pre-intermediate level in English. The images resulting from the subjects taking the WBT with the white background showed that a part of the frontopolar region in the frontal cortex near BA8 tended to experience higher levels of change in Hb concentration. These images suggested that the area for the frontal eye field, which roughly corresponds to BA8 [11], had been more highly activated than other brain areas while the subjects were taking the test with the white background. On the other hand, the regions which correspond to the areas related to language processing, including BA44, BA 45 and BA22 [12] tended to show higher Hb concentration changes while the subjects were taking the test regardless of background color.

5. Discussion and conclusion

In this study, the authors investigated if the performance of a Web-based English listening test taker differed, depending on the characteristics of a background color. We also examined how background colors can affect the functions of the test taker’s brain by observing relative changes in hemoglobin (Hb) concentrations in his/her brain by employing near-infrared spectroscopy. The brain function of each subject was studies in order to determine whether it can be affected by the background colors applied to computer screens, in particular the areas of the brain associated with linguistic tasks.
Fig. 4. Optical topography images showing relative changes in oxy Hb concentrations in the brain of a subject while taking the Web-based English listening test: (a) with the blue background; (b) with the white background

The average percentage of questions answered correctly by the subjects was a little higher for the blue background, but the test performance results did not show a significant difference between the blue and white backgrounds for the listening tests. A larger difference was found between the two colors for the previously conducted grammar tests. On the other hand, the NIRS results suggest that the background color can result in a difference in brain activity related to BA8, which is associated with eye movement activities. The results of this study indicate that blue may be slightly preferable to white for use as the background of a Web-based English listening test when question sentences are displayed using black text. The NIRS result also indicates that a white background provokes states that require more eye movements. The NIRS images of relative Hb concentration changes in subjects’ brains also suggest that the brain areas related to eye movements tended to be more active while performing the circle counting tasks with the white background than while doing the same task with the blue background. We suspect that a difference in test performance results between the listening test and the grammar tests may be due to the length of time the subjects were looking at the screen.

The questionnaire results for the concentration and fatigue levels of the subjects in this study indicate that the white background was better than the blue one for maintaining their concentration. This result contradicts our speculation that a white background may cause a decrease in test taker’s ability to concentrate. In Yamazaki’s study, the levels of fatigue and difficulty that test takers reported after taking the tests did not correlate with their test performance [8]. Other analyses in the study also showed no correlation between the subjects’ test scores and the values for the degree of fatigue that the subjects indicated. In this regard, the brain activity measured with NIRS should be further analyzed in relation to test taker’s concentration levels.

In this study, the authors did not employ a light blue background for observing blood hemoglobin concentrations in the brain for Web-based English tests. In Yamazaki’s study, a light blue background was associated with the highest average score for web-based and computer-based English grammar tests [8]. In the
next study, we plan to conduct experiments to see if Hb concentrations in Broca’s Area can be affected when the background color of WBTs is light blue. In addition, it is important to increase the number of subjects in the next study in order to obtain a more general sampling, since brain functions are known to vary among individuals.

Acknowledgements

The authors would like to extend our gratitude to Chieru Co., Ltd. for its assistance in developing a set of WBTs with different background colors. We also thank the students at Shibaura Institute of Technology who participated in this study. This work was supported by JSPS KAKENHI Grant Number 23501171.

References