The Evaluation of a Voting Web Based Application

Linda Harley, Keith Kline, Jerry Ray, Carrie Bell, Andrew Baranak, Chandler Price, Matthew Hung, and Brad Fain

Georgia Tech Research Institute, Atlanta, Georgia, USA
{linda.harley,keith.kline,jerry.ray,carrie.bell,andrew.baranak, chandler.price,matthew.hung,brad.fain}@gtri.gatech.edu

Abstract. Ballot layout and the incorporation of assistive technologies into voting systems are plagued with inconsistencies across the United States. The purpose of this study was to evaluate both ballot layout display configuration (information density) and a variety of controllers (e.g., mouse, 2- or 5-button controller) in order to assess performance and preference among voters. Participants were presented with three mock ballots, each with different layouts (scrollable pages, multiple columns or multiple pages per contest). Eye-tracking data and selection time data were recorded and a usability questionnaire was administered after each testing condition. The results of the study found that participants preferred the multiple column display configuration and the use of the mouse. The results from this study will be leveraged to design an iPad Voting Application with appropriate interfaces and controls. This will allow individuals with disabilities the opportunity to vote without requiring the dexterity to use a paper and pencil ballot.

Keywords: vote, web based application, human factors engineering, interface.

1 Introduction

The goal of the *Military Heroes Initiative and Military Voting* project is to better understand and improve voting technology and voting processes that affect recently injured military personnel who have returned from a combat zone with one or more disability. According to statistics compiled by the U.S. Department of Defense, as of February 15, 2013, 50,476 U.S. troops have been wounded during the conflicts in Iraq and Afghanistan, many returning home with a range of disabilities including loss of limbs, impaired vision, and traumatic brain injury [1]. These individuals may have difficulty getting to the polls to vote or even casting their vote via absentee ballot due to their functional and cognitive limitations.

To provide alternative solutions for absentee voting, GTRI developed a web-based voting application test-bed (Voting App), for use with smart technologies such as a smart phone or portable computer. This technology could facilitate obtaining an absentee ballot from an individual who is incapable of visiting the polling station. The purpose of the Voting App is to provide an alternative way of completing the absentee

M. Kurosu (Ed.): Human-Computer Interaction, Part III, HCII 2013, LNCS 8006, pp. 285–294, 2013.

ballot for those individuals who do not have the physical or mental capability to complete a standard paper and pencil absentee ballot. In the Voting App test-bed two variables were considered: (1) information density and (2) controllers.

1.1 Information Density

With an increase in digital reading, from reading on computers at work to leisurely eBook reading, finding an optimal presentation of information is important. To date, research investigating the optimal use of line length, multiple columns, and text justification is inconclusive. Longer line lengths result in faster reading times [2-5], but research suggests medium to short line length may result in greater comprehension [6-7]. Objective results support the use of both single columns of text [4], and multiple short columns [8] but subjective preference seem to be multiple short columns [4], [9-10].

Alternatively information can be displayed on multiple pages or on a scrollable page. Paging refers to navigating between pages where all text is replaced by a new screen full of information. Scrolling involves one page where text is replaced when new text appears on the screen. The objective is to ensure that individuals can move from page to page as efficiently as possible. If designers are unable to decide between paging and scrolling, it is recommended that they provide several short pages rather than one or two long pages that involve significant scrolling. However, with pages that have fast loading times, there may be no reliable difference between scrolling and paging when people are reading for comprehension [11]. The results of this study should help provide enough information to make confident decisions as to the appropriate layout of information.

1.2 Controllers

The most complete collections of hand anthropometry data have been collected by the US Army [12-13]. Studies have shown no significant difference in anthropometric hand measures between people with and without disabilities [14] or due to aging [15]. However, more longitudinal data are needed to fully confirm the aging effects. Where the considerable differences are generally seen is in the areas of strength and motor control, for example, with age there are considerable decreases in strength, dexterity, precision, coordination, joint mobility and sensitivity [16-17].

Push buttons have been shown to have a small space claim and are easy to operate [18]. According to recent occupation injury data, finger injuries are the most prevalent of upper extremity injuries [19]. While performance with controls such as knobs and dials are severely affected, push buttons show no performance difference between disabled and able-bodied users [20]. To better determine differences in button configurations two separate push button configurations were created for this test. The first of these configurations used two buttons, allowing for a *tab* and *select* control (Fig. 1. (A) 2-Button and (B) 5-Button configuration). The other configuration was a five button system, which allowed for *tab*, *shift* + *tab*, *left* (mapped to previous page command of the Voting App), *right* (mapped to the next page command of the Voting App), and *select* (Fig. 1. (A) 2-Button and (B) 5-Button configuration).

1.3 Hypotheses

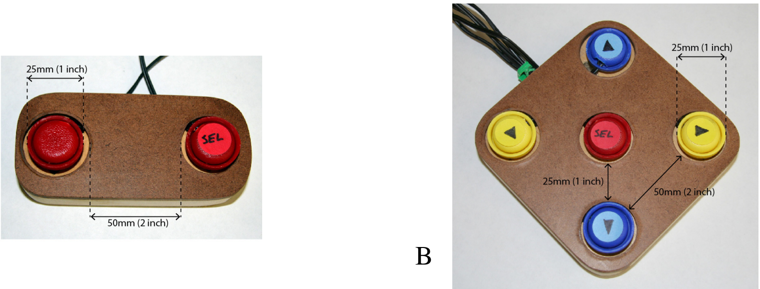
The purpose of this study was to examine information density and controllers for the Voting App.

Aim 1
Investigate the effect of different information density displays for the Voting App.

Hypothesis 1.1: The null hypothesis is that there is no significant difference between scrollable (*scroll*), multiple column (*columns*) or multiple page (*pages*) display configurations.

Aim 2
Investigate the effect of different controllers for the Voting App.

Hypothesis 2.1: The null hypothesis is that there is no significant difference between the *mouse*, *2-button* and *5-button* controls.



A B

Fig. 1. (A) 2-Button and (B) 5-Button configurations

2 Method

2.1 Participants

Eighteen (18) participants (26±13 years of age, 12 males and 6 females) volunteered for this study. Participants were 18 years of age or older and considered to be healthy. The Georgia Institute of Technology Internal Review Board approved the study.

2.2 Experimental Setup

Participants were asked to sit comfortably behind a desktop computer (Fig. 2. Participant using the 2-button controller to navigate the Voting App at the eye-tracking station.). The SmartEye system was used to track the participant's eye movements in order to determine areas of fixation on the screen. Three SmartEye cameras (Basler acA640-100gm cameras with 8mm lenses), each with two IR flasher devices, were used to track eye movement at a sampling rate of 60Hz. Three controllers were used in this study: (1) *mouse*, (2) *2-button* controller (Fig. 1. (A) 2-Button and (B) 5-Button configuration) and (3) *5-button* controller (Fig. 1. (A) 2-Button and (B) 5-Button configuration). The buttons (Enabling Devices Compact Switch #745) were connected to a switch interface device (X-keys XSI-38-US). The controllers interfaced the 3.5 mm switch plugs to the PC via a USB port, and facilitated programming the switches with the desired keyboard inputs. These buttons were mapped to standard keyboard keys, which were then encoded in the Voting App.

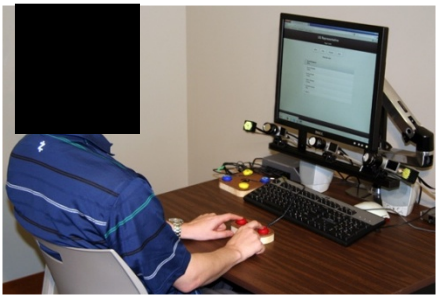


Fig. 2. Participant using the 2-button controller to navigate the Voting App at the eye-tracking station

2.3 Experimental Groups

Participants were pseudo-randomly assigned to one of three groups so that age and gender were balanced between groups. This study used a 3x3 factorial design, with the different controllers as a between-subjects group variable and the different display configurations as a within-subject variable. Each participant, assigned to only one of the three controllers, was asked to complete the mock ballot under each of the three display configuration conditions. The order in which the display configurations were tested was counterbalanced across participants by Latin square to eliminate order effects.

2.4 Task Protocol

After consent had been obtained, participants were asked to sit at the SmartEye computer while the cameras were calibrated to their individual anthropometric characteristics. After calibration, the participants were instructed on how to use the controller and presented with one of the ballot configurations. Participants were instructed to experiment with the ballot and take their time. After completion of a ballot, participants completed the System Usability Scale questionnaire. Then participants were given a few minutes to rest, while the SmartEye system saved the data. The process was then repeated using a different ballot configuration. At the completion of the study, participants were asked several interview questions to determine any additional difficulties they may have encountered.

2.5 Data Analysis

Time Dependent and User Selection Variables

The Voting App contained a built-in function that logged each keystroke and selection made by participants along with the time stamp. The following variables were calculated:

- Ballot Duration: The total time it took the participant to complete the ballot. Time started when the participant clicked on the "Start" button and stopped when the participant clicked the "Submit" button.
- Mean Click Time: Click time is defined as the time between selections made by the participant. Click time was averaged across the entire ballot to calculate Mean Click Time.
- Overvote: An overvote occurred when the participant attempted to vote for more participants than they were allowed to for that particular contest. The number of overvote's that occurred for the entire ballot was tabulated.
- Undervote: An undervote occurred when the participant attempted to vote for less candidates than they were allowed to for that particular contest. The number of undervotes' that occurred for the entire ballot was tabulated.
- Back Button: The number of times the "Back" button was selected (Fig. 3. Screenshot of ballot layout).
- Help Button: The number of times the "Help" button was selected (Fig. 3. Screenshot of ballot layout).
- Review Button: number of times the "Review" button was selected. This button allowed participants to review their entire ballot (Fig. 3. Screenshot of ballot layout).
- Next Button: The number of times the "Next" button was selected. This button allowed participants to go to the next page (Fig. 3. Screenshot of ballot layout).

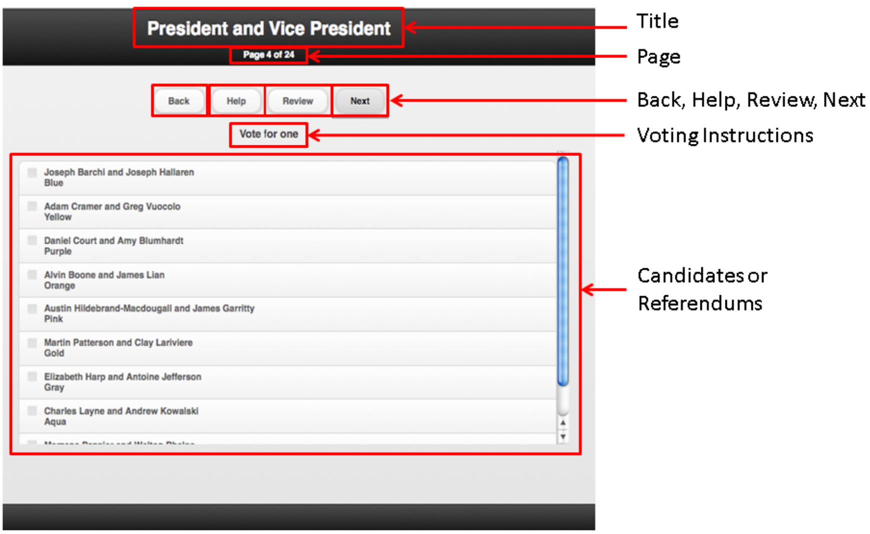


Fig. 3. Screenshot of ballot layout

Eye Tracker Dependent Variables

The eye tracker system was used to determine how much time participants spent viewing several regions of interest on the Voting App. The following variables were recorded and analyzed using the SmartEye software. The look zones for the study began recording data when the participant was presented with the first page and stopped recording once the user submitted the ballot. The time spent viewing each of the following regions of interest were monitored (Fig. 3. Screenshot of ballot layout): (1) Candidates and referendums, (2) Title, (3) Voting Instructions, (4) Page, (5) Back, (6) Help, (7) Review, (8) Next.

3 Results

3.1 Time Dependent and User Selection Variables

A two-way ANOVA was run with display configuration (*pages*, *column* and *scroll*) and controls (*5-button*, *2-button* and *mouse*) being the independent variables. The statistical analysis found no significance between the independent variables for the ballot duration, mean click time, number of undervotes, or number of times that the "Help" button was selected ($p>.05$ for all comparisons). However, a significantly greater number of overvotes occurred for the *pages* display configuration when compared to both the *column* and *scroll* display configurations, $F(2,30) = 8.73$, $p = 0.001$. This is supported by the fact that the number of times the "Back" and "Next" buttons were selected was significantly greater for the *pages* display configuration when compared to that of the *column* and *scroll* configurations, $F(2,30) = 80.78$, $p<0.001$. For the controllers, the only significance among conditions was that the "Review" button was selected more frequently for the *2-button* controller than the *5-button* and *mouse* controllers, $F(2,15) = 5.98$, $p = 0.012$. There were no significant interactions between display configurations and controls.

When considering the difference between display configurations, these results suggest that many more errors occurred for the *pages* display configuration than the *columns* and *scroll* configurations. This resulted in a greater amount of overvotes occurring, meaning an increased navigation between *pages* as far as the number of times that the "Next" and "Back" buttons were selected. One possible explanation may be that cognitive load was high for the *pages* display configuration, since participants were using working memory to recollect how many candidates they had selected for the contest. When evaluating the difference between controllers, the only result of note was that participants in the *2-button* condition more frequently selected the "Review" button. This strategy may have assisted participants in avoiding unnecessary navigation between the pages, as the *2-button* controller only allowed for forward linear advancement and selection. Although not significant, participants using the *mouse* had the shortest ballot duration and mean click times. These results indicate that the most optimal controller was the *mouse* and that the *pages* display configuration was not an optimal display configuration. There were no significant differences between *scroll* and *column* display configuration.

3.2 Eye Tracker Dependent Variables

A two-way ANOVA was run with display configuration (*pages*, *column* and *scroll*) and controls (*5-button*, *2-button* and *mouse*) as the independent variables, and time spent viewing regions of interest on the voting app as the dependent variable. The statistical analysis found no significance difference between the independent variables for the time spent viewing the "Candidates and referendum", "Voting instructions", "Page numbers", "Back", "Help", and "Next" buttons. However, participants spent a significantly greater amount of time viewing the "Title" for the *pages* display configuration when compared to the *column* and *scroll* display configurations, $F(2,20) = 10.99$, $p < 0.01$. In addition, results showed that participants spent more time viewing the "Title" for the *5-button* controls than the *2-button* controls. In fact, the time participants viewed the "Title" for the *2-button* controls was small and negligible when compared to the other conditions. This may be due to the linear nature of the navigation.

These results support that of the previous section, in that participants in the *pages* display configuration had to pay much closer attention to the title of the page to understand which contest they were on.

3.3 Subjective Rating Results

System Usability Scale

A 2-way analysis of variance was conducted on the System Usability Scale data, with display configuration as a within-subjects variable and controls as a between-subjects variable. The data are shown in Fig. 4. Subjective ease of use ratings obtained with

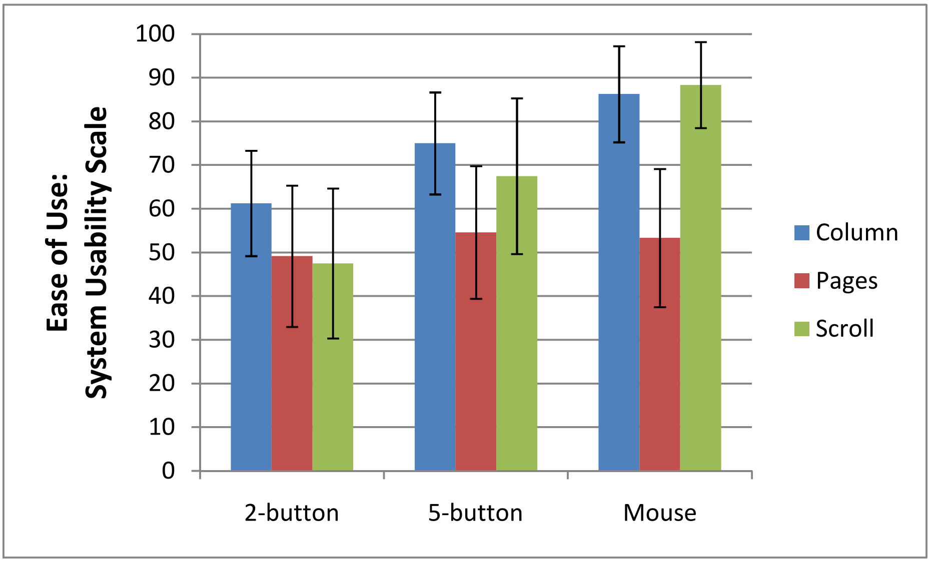


Fig. 4. Subjective ease of use ratings obtained with the System Usability Scale. Error bars represent 95% confidence intervals.

the System Usability Scale. Error bars represent 95% confidence intervals The effect of display configuration was significant, $F(2,30) = 7.99$, $p = 0.002$. The effect of controls was not significant, $F(2,15) = 3.15$, $p = 0.07$. The interaction between controls and display configuration was not significant $F(4,30) = 1.82$, $p = 0.09$.

Post-hoc paired t-tests were conducted on the display configuration conditions. The critical p-value was adjusted to 0.0167 for multiple comparisons. Ratings for the *column* condition were higher than those in the *pages* condition, $t(17) = 4.95$, $p<0.01$, but ratings for the *column* condition were not significantly higher than those in the *scroll* condition, $t(17) = 1.02$, $p = 0.32$. Ratings for the *scroll* condition were not significantly higher than for the *pages* condition, $t(17) = 2.30$, $p = 0.03$.

4 Discussion

The purpose of this research was to examine different information density display configurations and different controllers for the Voting Application. The first null hypothesis was that there would be no significant difference between the *scroll*, *columns* or *pages* display configurations. The results of this study suggested that there was no significant differences between display configurations when considering the length of time it took participants to complete the ballot or the mean time between selections. However, the number of overvotes that occurred during the *pages* display configuration was significantly higher than the *columns* or *scroll* display configurations. Participants also selected the “Back” and “Next” navigation buttons a significantly greater amount of times for the *pages* display configuration when compared to the other conditions. This implies that perhaps the cognitive load on recalling vote selections as well as understanding what contest users were viewing was too complex. This is supported by the fact that participants spent an increased amount of time looking at the “Title” of the contests for the *pages* display configuration than the other conditions. Moreover, results from the System Usability Scale showed that participants had difficulty with the *pages* display configuration. Participants reported that the *pages* display configuration was by far the most difficult, confusing and memory intensive. There were no significant differences in performance of participants between the *scroll* and *column* display configurations. Therefore, the research concludes that the *pages* display configuration is the least optimal solution for presenting information on a ballot, and that either *scroll* or *column* display configurations should be used.

The second null hypothesis was that there would be no significant difference between the *mouse*, *2-button*, and *5-button* controls. Participants utilized the “Review” button more often in the *2-button* condition when compared to the *5-button* or *mouse* conditions. This approach may have been useful in avoiding cycling through all the contests to select only those contests in which they were interested. Recall that the *2-button* controls only allowed for one-directional (tab-forward) movement throughout a page. Furthermore, participants spent little time looking at the “Title” for the *2-button* control (492±225 milliseconds). If participants navigated to the contest from the review page, they would not need to read the title, as they would already know what contest they were going to vote on. Participants reported that the *2-button* control was cumbersome. The *mouse* control was the easiest to use, which is likely due to the fact that participants are already familiar with this technology. These results lead us to conclude that the *2-button* control is the least optimal solution, while the *5-button* control is mediocre and the *mouse* is preferred. A more optimal button solution may be 3-button controller, since it will allow forward and backward navigation as well as a selection button. This would be less demanding than the *5-button* controller and more flexible than *2-button* controller.

One of the main goals of the Voting App test-bed is to utilize it to evaluate a variety of long-standing issues being discussed in the voting world, such as the use of plain language in instructional content, font types, cognitive load, symbolism, and so forth. In addition to the Voting App test-bed, research staff is also in the process of designing an iPad hard case with a 3-button interface to be used in testing.

References

1. U.S. Casualty Status for OI, OND, and OEF, `http://www.defense.gov/news/casualty.pdf`
2. Duchnicky, R.L., Kolers, P.A.: Readability of text scrolled on visual display terminals as a function of window size. Human Factors 25, 683–692 (1983)
3. Dyson, M.C., Kipping, G.J.: The legibility of screen formats: Are three columns better than one? Computers and Graphics 21(6), 703–712 (1997)
4. Dyson, M.C., Kipping, G.J.: The effects of line length and method of movement on patterns of reading from screen. Visible Language 32, 150–181 (1998)
5. Dyson, M.C., Haselgrove, M.: The influence of reading speed and line length on the effectiveness of reading from screen. Int. J. Human-Computer Studies 54, 585–612 (2001)
6. De Brujin, D., De Mul, S., Van oostendorp, H.: The influence of screen size and text layout on the study of text. Behavior and Information Technology 11, 71–78 (1992)
7. Chaparro, B., Baker, J.R., Shaikh, A.D., Hull, S., Brady, L.: Reading online text: A comparison four whitespace layouts. Usability News 6(2) (2004)
8. Lam, K., Lam, Y., Liu, J., Shin, U.G.: Reading comprehension and rate: one column versus three columns. Dept of Computer Science Class Project, Univ. of Maryland, College Park (2000)
9. Bouma, H.: Visual reading processes and the quality of text displays. In: Grandjean, E., Vigliani, E. (eds.) Ergonomic Aspects of Visual Display Terminals. Taylor & Francis, London (1980)
10. Andrevey, V., Martynov, A.: Effects of splitting text into multiple columns. Dept of Computer Science class Project, Unv. Of Maryland, College Park (2000)
11. U.S. Department of Health and Human Services: The research based web design and usability guidelines (2006)
12. White, R.M.: Comparative anthropometry of the hand. U.S. Army Natick Research & Development Laboratories (1980)
13. Grenier, T.M.: Hand anthropometry of U.S. Army Personnel. U.S. Army Natic Research, Development, and Engineering Center (1991)
14. Feeny, R.: Specific anthropometric and strength data of people with dexterity disability. In: Consumer and Competition Policy Directorate: Department of Trade and Industry (2002)
15. Molenbroek, J.F.M.: Anthropometry of elderly people in the Netherlands; research and applications. Applied Ergonomics 18(3), 187–199 (1987)

16. Daams, B.J.: Human force exertion in user-product interaction. Doctoral Thesis, TU Delft, Amsterdam (1994)
17. Haigh, R.: The ageing process: a challenge for design. Applied Ergonomics 24(1), 9–14 (1993)
18. Van Cott, H.P., Kinkade, R.G.: Human engineering guide to equipment design. United States Dept. of Defense, Joint Service Steering Committee (1972)
19. Bureau of Labor Statistics: Nonfatal occupational injuries and illnesses requiring days away from work. 2010 US Department of Labor (2011)
20. Pennathur, A., Mital, A., Contreras, L.R.: Performance reduction in finger amputees when reaching and operating common control devices: A pilot experimental investigation using a simulated finger disability. Journal of Occupational Rehabilitation 11(4), 281–290 (2001)