The Information Technology and Innovation Foundation

Accessible Voting Technology Initiative

Working Paper Series

Working Paper #012

A Web Based Voting Application Study of Display Layouts for Older Adult Voters with Arthritis

Linda Harley, Keith Kline, Chandler Price, Adrienne Jones, Shaun Mosley, Sarah Farmer and Brad Fain

December 2013

The Information Technology and Innovation Foundation
1101 K Street NW, Suite 610
Washington, DC 20005
(202) 449-1351

This research was supported by the U.S. Election Assistance Commission (EAC) under grant number EAC110149B. Any opinions, findings, conclusions or recommendations expressed in this report are those of the authors and do not necessarily represent the views of EAC or ITIF.

Preface

This report is part of a series of working papers. Some of the content in the introduction and methods of this report are identical to those of the other reports in the series, because the research studies had similar justifications and methods.

Executive Summary

The Help America Vote Act (HAVA) legislation was passed by Congress in 2002 in response to the controversy surrounding the 2000 U.S. presidential election. As a result many states changed their procedures and equipment to be more accessible and usable to voters who are hearing or visually impaired or who use a wheel-chair. However, there are many hidden disabilities that are not as apparent that should also be considered, such as functional limitations due to arthritis, cognitive deficits, or upper mobility impairments. The CDC estimates that approximately 50 million adults in the U.S. have self-reported being diagnosed with arthritis. Arthritis limits the functional ability of voters that may directly impact their performance in using a voting system. For example, voters with arthritis may experience difficulty pressing buttons repetitively and therefore find it cumbersome to navigate ballots. The study of how older adults with functional limitations, such as arthritis, interact with voting systems can lead to better recommendations for universally accessible voting systems.

The goal of this research effort was to enable private and independent voting by older adult voters with arthritis. This research study **investigated the usability of three different display layouts**: multi-page, multi-column, and scrolling layouts. The display layouts may require more or less tactile button presses, which to an individual with arthritis is a consideration as it affects the comfort level in using the voting system. The multi-page layout refers to presenting all candidates' names for a single contest across multiple pages of the ballot. The multi-column layout displays all candidates simultaneously on a single page listing the names in multiple columns. The scrolling layout involves one column where the text is replaced with new text as the participant scrolls up and down the page.

Twelve older adult participants with arthritis in the hands, wrists, and/or fingers volunteered for this study. Participants interacted with each ballot displayed in each layout in a within participant design. The order in which the display layouts were presented to participants was counterbalanced. Participants used the 3-button input device to navigate through a web based voting application (Voting App) while their performances were monitored. The Voting App produced an event log that logged the time at which various events occurred, such as selecting a candidate. Eye tracking data was captured to determine where participants were looking during each task. After using each display layout, participants were asked a series of questions. At the end participants were asked to rank the overall preference of the display layouts.

Participants ranked the multi-column layout more favorably than the multi-page or scrolling layout, and the multi-page layout received the least-favorable ranking of all the display layouts, although these differences were not statistically significant. The multi-page layout took participants the longest time to complete, resulted in the most over-votes, and took participants significantly longer to find the targeted candidate. There was relatively little difference in performance between the multi-column and scrolling layouts. Some participants indicated that they preferred the multi-column because they were able to see all of the candidates simultaneously. However, for a long list of candidate names the multi-column layout approach is often not feasible due to the limited screen space and the need to keep the names at a decent font size to promote readability. It is recommended in these instances that the information be shown in a scrolling layout.

Table of Contents

List of Figures

List of Tables

Introduction

In 2002, the Help America Vote Act (HAVA) was passed by congress in response to the controversy surrounding the 2000 U.S. presidential election. HAVA required that each polling location have at least one voting system that is accessible to individuals with disabilities. The implementation of this has been to ensure that the voting systems are accessible to voters who are hearing or visually impaired, or who use a wheelchair. However, there are many hidden disabilities that to date have not been addressed adequately, such as arthritis.

Lawrence et al (1998) estimate that 59.4 million Americans will be affected by Arthritis by the year 2020. According to the CDC, arthritis is the most common cause of disability among U.S. adults. The CDC analyzed data from the 2010-2012 National Health Interview Survey (NHIS) and found that 52.5 million adults (22.7% of total respondents) self-reported that they had been diagnosed with arthritis; 22.7 million (9.8% of total respondents) reported that arthritis had attributed to the limitations in their daily activities (MMWR 2013). The highest prevalence of arthritis was in those over the age of 45, and these individuals often had comorbidity such as heart disease, diabetes, and fatigue. The older adult population is of particular interest because conditions such as osteoarthritis are often associated with age. In addition, older adults tend to have lower working memory capacity than younger adults (Hambrick & Engle, 2012) and tend to learn procedures less easily than younger adults (Harada et al., 2010).

Individuals with arthritis have functional limitations that include difficulty reaching, grasping and pushing objects. Alvarez and Hall (2010, as cited in Hall, 2013) found that tasks such as putting a ballot in a ballot box or holding a pen, and other physical tasks associated with a physical ballot can be difficult or impossible for some voters with physical disabilities. Arthritis in the fingers may limit their range of motion in those joints, and result in pain if they have to perform repetitive tasks, such as pressing buttons on a voting system. This could also lead to difficulty sliding the voter card to activate the voting system, because this task requires some degree of precision and dexterity.

The purpose of the current study was to determine which display layout would be easiest to use with a 3-button input device. The display layouts may require more or less tactile button presses, which to an individual with arthritis is a consideration as it affects the comfort level in using the voting system.

Voting Application

GTRI developed a web based voting application test bed, henceforth referred to as the “Voting App”, to be used for testing user interface concepts for a variety of applications and to be used with technologies such as portable computers (Harley et al., 2013). The Voting App is designed to be accessible to individuals with impairments in vision, hearing, upper or lower mobility and cognitive deficits. The Voting App provides designers a means of testing user interface design alternatives with individuals with disabilities. The Voting App captures objective measures of performance, including selection times, ballot duration, sequence of selections, and errors. These objective data, along with subjective usability questionnaires and rating scales, can help designers of voting systems understand how individuals with disabilities interact with their product.

Eye Tracking

While the Voting App recorded overt user interactions with the system, an eye tracking system was used to track participants' gazes. This data was expected to provide additional information about the tasks that users were performing, such as visual search and attention to the on-screen navigation controls.

Display Layouts

With an increase in digital reading mediums, finding the optimal way to represent information is vital to the ease of use and usability of such systems. Information can be displayed on multi- page, multi-column, or scrolling layout. Scrolling layout involves one column where the text is replaced with new text as the user scrolls up and down the page. Multi-page layout refers to ballots in which candidates in a single race may span multiple pages. The multi-column layout displays all the candidates on a single screen and multiple columns.

The research on optimal line length, multiple columns, and text justifications remains inconclusive (Duchnicky & Kolers, 1983; Dyson & Kipping, 1997, 1998; Dyson & Haselgrove, 2001). Some studies support use of single columns of text, whereas others argue that multiple short columns are easier to read (Andrevey & Martynov, 2000; Bouma, 1980; Chaparro et. al., 2004; De Brujin et. al., 1992; Lam et. al., 2000).

The aforementioned studies focused on reading sentences or paragraphs of text, which is directly relevant to the display of amendments on ballots, but might not be as applicable to the display of candidate names. The studies did not address the navigation in the various layouts, such as moving an on-screen cursor or finding a candidate, which are crucial tasks for users of electronic ballots.

Hypotheses

The present study examined three design alternatives for ballot layouts that differed with respect to how information was displayed on high-content races, in which the number of candidates or length of an amendment was too long to fit in a single column on a single page. The three display layouts included a scrolling layout (Figure **1**.a), a multi-column layout (Figure **1**.b), and a multi-page layout (Figure **1**.c).

The multi-page layout theoretically would enable users to make selections with fewer button presses, because they could move between batches (pages) of candidates with a single button press instead of moving the cursor through the list of candidates. However, the multi-page layout – as well as the scrolling layout – prevent users from seeing all candidate names at once. It may be easier for users to visually scan the full list of candidates presented at one time in a multi-column layout. Thus, there might be a trade-off between efficiency of button presses, which might favor the multi-page layout, and ease of visual scanning, which might favor the multi-column layout.

H1. Subjective user preferences and objective measures of performance were expected to be better for the multi-column layout than for the scrolling layout and multi-page layout.

H2. The multi-column layout enables users to see all of the candidate names simultaneously, so that they can see how many candidates they have voted for on a given race. The multi-column layout may also make it easier for users to visually scan for a desired name. Therefore, gaze times on the candidates are expected to be shorter for the multi-column layout.

H3. The multi-page layout provides a possible benefit of more efficient use: Users can navigate between groups (pages) of candidates with a single button click, rather than moving the cursor through all of the candidates on a race.

> **H3a.** However, some users might not recognize this benefit, and might be confused if they fail to realize that a single race can span multiple pages.
>
> **H3b.** The multi-page layout was expected to lead to more over-vote errors than the other layouts, because users might sometimes think that they have moved to another race rather than another page within the same race.

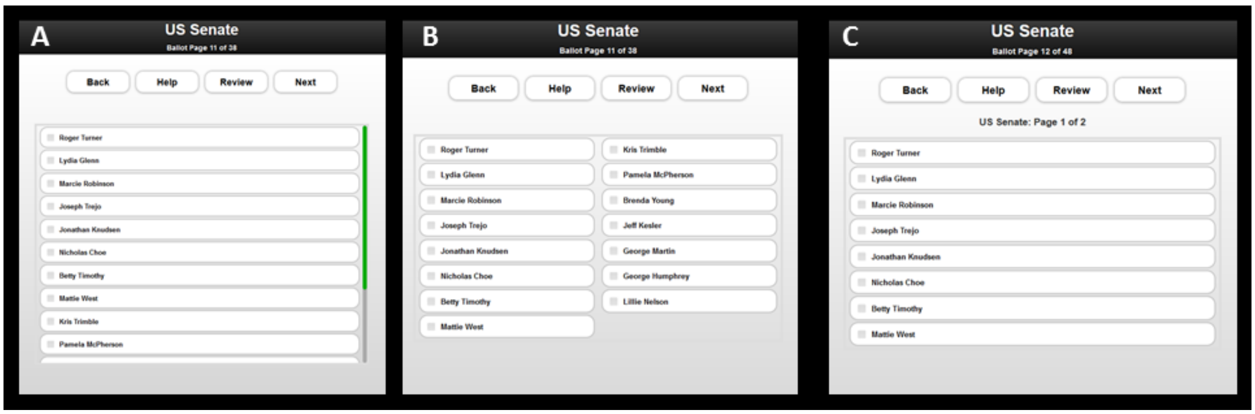

Figure 1. Different display layouts: (a) Scrolling layout, in which candidates were listed in a scrolling window for a given race, (b) Multi-column layout, in which candidates spanned multiple columns for a given race, (c) Multi-page layout, in which candidates spanned multiple pages for a given race.

Method

Participants

Twelve participants (mean age = 70.3±7.6; range = 56-81, 11 female, 1 male) volunteered for this study. In accordance with inclusion criteria, all participants reported symptoms of arthritis in the wrists, hands, or fingers. Participants were compensated $30 for their participation in the 1.5 hour experiment. The Georgia Institute of Technology Institutional Review Board approved the study.

Ballots

Each ballot consisted of 16 items including 12 races, a constitutional amendment, a ballot initiative, and two retention questions (e.g., "Shall John Doe be kept as city comptroller?"). These were preceded by instructions and a practice race. The final page of the ballot was a review page that displayed the participant's selections for all of the races.

Each page of the ballot included four on-screen control buttons (Back, Help, Review, Next; Figure 1). For races with more than eight candidates, the candidates were split across multi-column, multi-page, or scrolling layout as needed. Prior to each race participants received an on-screen prompt to vote for a particular candidate in the next race. Participants received pop-up warnings if they under-voted or over-voted on any race. The ballot review page also gave a visual indication of under-votes.

The scrolling layout had an indicator bar to the right showing participants where they were with respect to the entire list. In addition, the top or bottom half of the candidate names at the top and bottom of the list, respectively, were clipped in order to visually indicate that more information was available above/below the current location in the list (Figure 1.a). The multi-column layout consisted of two columns of candidate names, all displayed simultaneously on a single page (Figure 1.b). The multi-page layout consisted of showing a maximum of 8 candidates per page and each page was replaced by another page requiring participants to skip to next page to see the next set of 8 candidates (Figure 1.c).

Three equivalent ballots were created for use with *each* of the three layouts. The order in which the layouts/ballots were used was counterbalanced by Graeco-Latin square. The three ballots were matched with regards to the list-locations of the correct (prompted) candidates in the races. For example, a prompted candidate would appear as number 8 in a list of 15 candidates for a given race on all three ballots. The layout of the lists would vary across the ballots (multi-page, multi-columns, scrolling), but the prompted candidate would always appear in list position 8 for that race.

User Input Device

Participants used a 3-button input device to navigate the ballot (Figure 2). The 3-button device featured a select button, a cursor-forward button, and a cursor-backward button. The latter two were used to move the on-screen cursor through the on-screen options. When the desired item was highlighted, the user pressed the select button to actuate it. When the cursor reached the end of the page, the cursor-forward button caused it to wrap back to the top of the page. The buttons (Enabling Devices Compact Switch #745) were connected to a switch device (X-keys XSI-38-US). The switch device was connected by

3.5 mm switch plugs of the buttons to the PC via a USB port. The switch inputs were mapped onto the desired keyboard inputs to control the Voting App.

Figure 2. 3-button input device.

Eye Tracking System

A SmartEye eye tracking system was used to track participants' direction of gaze as they voted. Three SmartEye cameras (Basler acA640-100gm cameras with 8mm lenses) with two IR flasher devices were used to sample eye position at 60Hz (Figure 3).

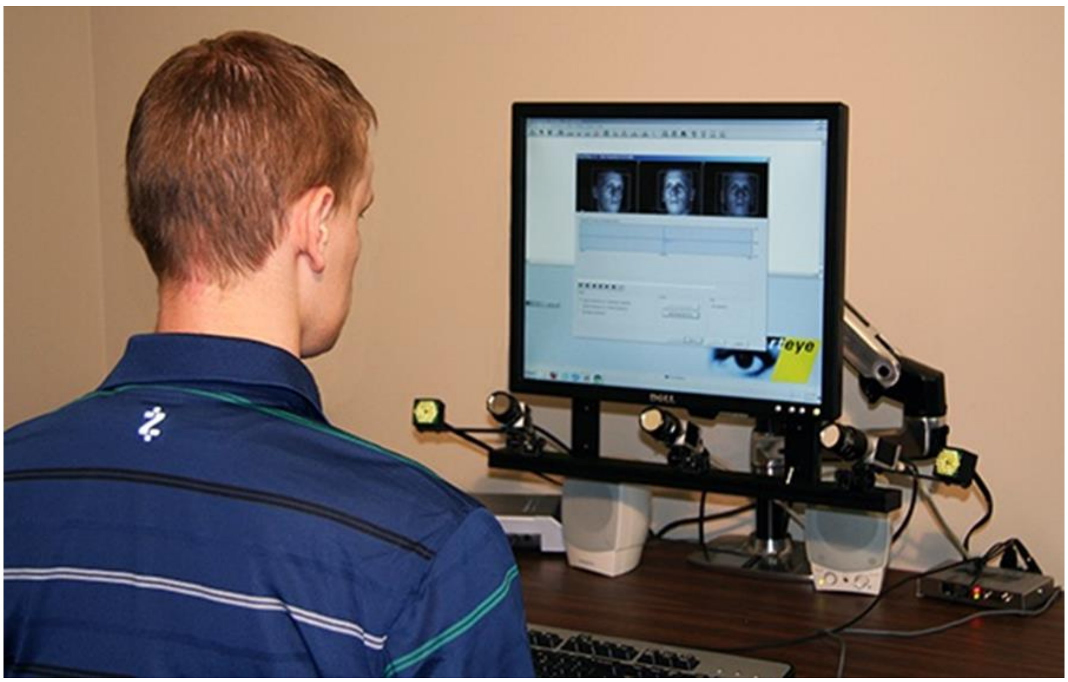

Figure 3. SmartEye tracking system.

Design

A repeated measures design was used. Each participant interacted with each display layout. Three equivalent ballots were created for use with each of the three layouts.

Display layout (scrolling, multi-column, and multi-page) was the independent variable for all analyses of user performance, user preference, and eye tracking data. Repeated measures ANOVAs were conducted on each dependent variable.

Subjective dependent variables included ratings of pain, System Usability Scale (SUS), Voting System Usability Scale (VSUS), and display layout ranking. Objective dependent variables were derived from event logs (user selections) and eye tracking data.

Event logging variables included the following:

- *Over-votes*: The number of races in which a participant over-voted.
- *Under-votes*: The number of races in which a participant under-voted.
- *Percent correct*: Percentage of candidate selections that matched the prompts.
- *Click Time*: The time elapsed between selections of any on-screen buttons, including navigation control buttons and candidates' names (seconds).
- *Response Time*: The time elapsed between the onset of a race and the participant's selection of a candidate divided by the position of the candidate in the list (seconds/candidate).
- *Race Time*: The average time spent on the four races on each ballot in which the candidate list was sufficiently long to require presentation across multi-column, multi-page, or scrolling layout (seconds).

Dependent variables for the eye tracking analysis included the following:

- *Controls Gaze Time*: The amount of time spent looking at the on-screen control buttons (i.e., back, help, review, next; seconds).
- *Candidates Gaze Time*: The amount of time spent looking at the entire list of candidates (seconds).
- *Target Candidate Gaze Time*: The amount of time spent looking at the target candidate (seconds). This is a subset of *Candidates Gaze Time.*
- *Visual Search Time*: The amount of time spent looking at all the options except the target candidate (seconds). This is the difference in time between the *Candidates Gaze Time* and the *Target Candidate Gaze Time*, and reflects the amount of time participants spent seeking the target.

Procedure

Participants completed the experiment in one-on-one sessions with an Experimenter, who was present throughout the session. Participants first signed an informed consent form and received a brief overview of the procedures from the experimenter. Then participants completed a short questionnaire on their arthritis symptoms and voting experience, and the participant's pinch strength and grip strength were measured. Grip strength assessment was performed with a Jamar Hydraulic Hand Dynamometer (5030J1 Serial #30303257). Users were instructed to grip the dynamometer as tightly as possible without experiencing excessive discomfort. Both left and right hands were tested. The pinch strength measurement was performed with a Jamar Hydraulic Pinch Gauge (7498-05 Serial # 60203139). Three pinch types (Tip, Key, and Palmar) were performed. For the tip pinch, users placed the tips of their index finger and thumb on opposing sides of the gauge. For the key pinch, users placed the lateral aspect of the middle phalanx of their index finger on the bottom of the gauge and their thumb on the top of the gauge. For the palmar pinch, users placed their thumb on the top of the gauge and the pads of their index and middle fingers on the bottom of the gauge.

Next, participants were seated comfortably in front of the computer and eye tracking system. The experimenter calibrated the eye tracking system and then launched the Voting App. The participant completed the first ballot while the experimenter observed. Immediately afterward, the participant completed a set of ratings scales for the display layout, which included a pain rating scale, the System Usability Scale (Sauro, 2011), and a usability scale created specifically for the Voting App. The procedures were repeated for the next two ballots.

After all three layouts had been presented, the experimenter interviewed the participant. Each participant ranked the layouts based on their subjective preference. Participants explained their rankings and provided any additional comments. To conclude the session, participants were debriefed and compensated for their time.

Results

Arthritis Symptoms

All participants reported experiencing arthritis symptoms in their fingers, hands, or wrists on the day of the study. Figure 4 shows the reported distribution and severity of arthritis symptoms. Eight participants reported having only osteoarthritis/degenerative arthritis. One participant reported having osteoarthritis and ankylosing spondylitis. One participant reported having osteoarthritis, rheumatoid arthritis, and systematic lupus erythematosis. One participant reported having only rheumatoid arthritis. One participant reported having osteoarthritis, ankylosing spondylitis, and psoriatic arthritis.

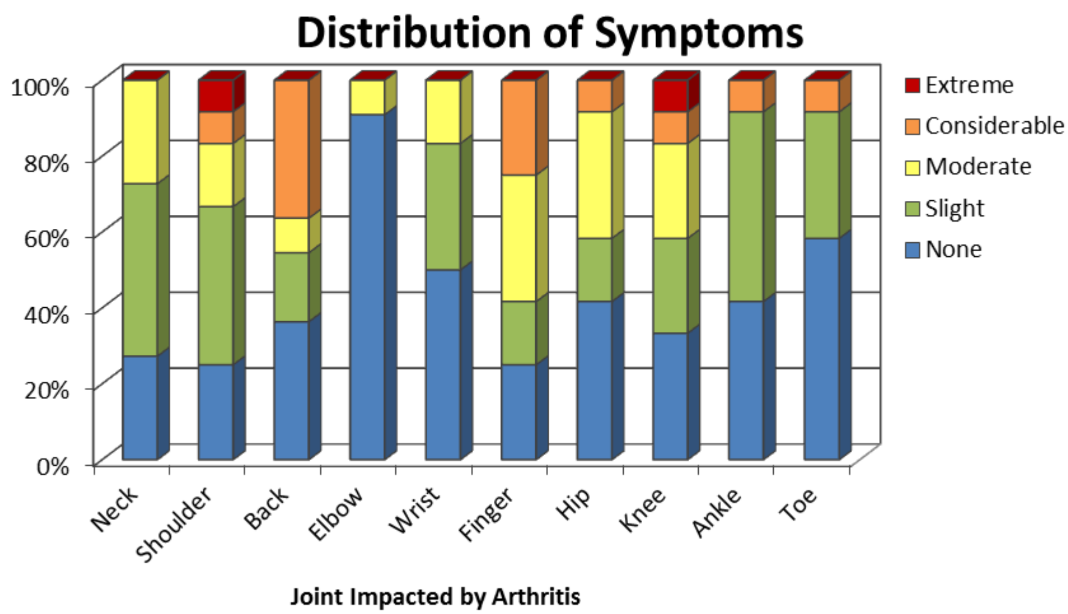

Figure 4. Distribution of arthritis symptoms.

Participants were asked to rate the arthritis-related difficulty that they had experienced using touchscreen devices, buttons, and voting systems (Table 1). Three participants reported having slight difficulty to moderate difficulty using a touch screen or pressing buttons due to their arthritis. Only two participants reported having slight difficulty voting at a polling place due to arthritis. The majority of participants experienced no difficulty in performing these tasks.

Table 1. Arthritis user experience questionnaire.

Question	No Difficulty	Slight Difficulty	Moderate Difficulty	Extreme Difficulty
Do you have difficulty using a touch screen due to your arthritis?	9	3	0	0
Do you have difficulty pressing buttons due to your arthritis?	9	2	1	0
Did you have any difficulty voting at a polling place due to your arthritis?	10	2	0	0

Strength Assessment

Grip and pinch strength are related to the use of the tactile interface. It was important to characterize participants' hand strengths to identify any outliers (extremely strong or weak participants). Both the left and right hands were tested. Figure 5 contains the results of the average grip strength assessment using a wide (open) grip and a narrow (closed) grip. Figure 6 contains the averaged results of the pinch force assessment for the tip, key, and palmar pinch. In both figures, the mean force is represented by the filled diamond, median is represented by horizontal line, interquartile range (25^{th} to 75^{th} percentile) is the shaded box, plus or minus 1.5 interquartile range is represented by the vertical lines, and outliers are unfilled squares. None of the grip or pinch strength measurements showed any participants as outliers, which were defined as measurements three standard deviations above or below the mean.

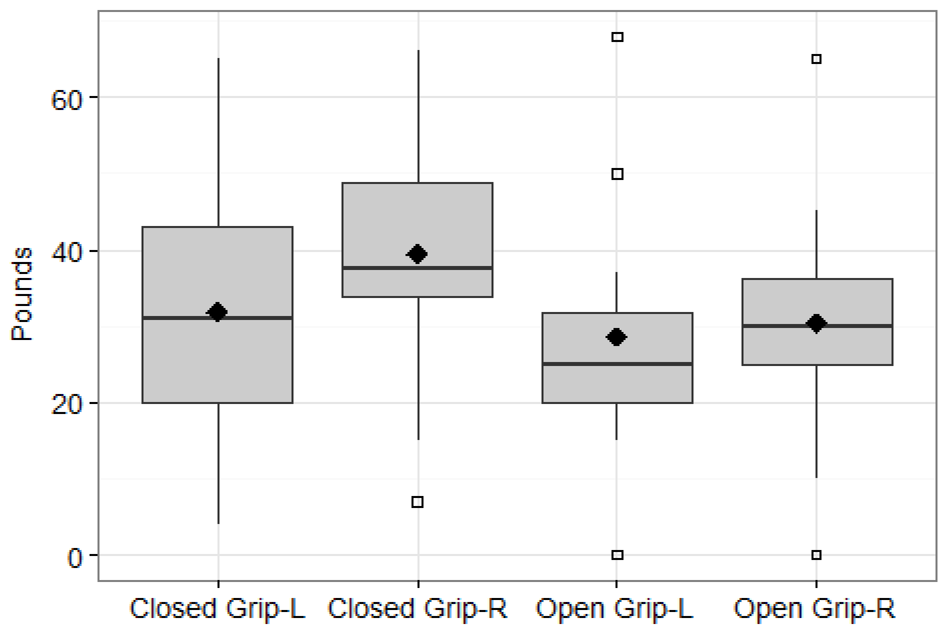

Figure 5. Grip strength measurements.

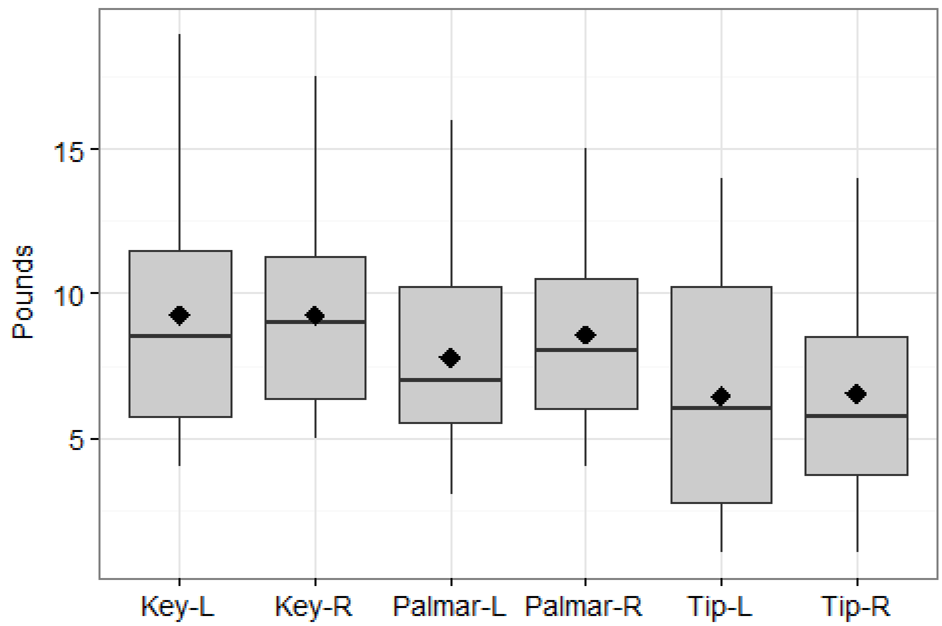

Figure 6. Pinch strength measurements.

Subjective Data

Rankings

The multi-column layout was expected to receive higher user rankings (H1), but a Friedman rank test did not show significant differences in rankings among the three display layouts, $X^2(2) = 3.5$, $p = .17$. However, the multi-column layout was most frequently ranked the highest (Table 2). It also received the highest total score and ratio of favorite to least favorite votes. "Total score" was computed for each layout by assigning scores of 3, 2, and 1 for favorite, second favorite, and least favorite, respectively, and summing across participants. Ratios of favorite-to-least-favorites were computed by dividing the respective cells in the table. Based on both Total Scores and the Ratios, there was a slight but statistically insignificant preference for the multi-column layout.

Table 2. Subjective rankings of preference for the three display layouts. Values in the cells represent counts.

	Multi-Column	Multi-Page	Scrolling
Favorite	7	2	3
Second Favorite	2	3	7
Least Favorite	3	7	2
Total Score	28	19	25
Ratio Fav/Least Fav	2.33	0.29	1.50

Note: Higher "Total Score" and "Ratio Fav/Least Fav" represent higher desirability.

Seven of the twelve participants reported that the multi-page layout was their least favorite; five of these seven reported feeling lost or not understanding at first that the candidate list spanned multiple pages. Seven participants reported that the multi-column layout was their favorite; five of these reported that they liked being able to see all of the candidates at one time. It was easier to visually scan the columns for the name they were seeking. Participants who preferred the scrolling layout stated that they like to see all of the candidates "in one place" or "all aligned in one column."

Usability Ratings

Participants also completed rating scales for each input device. Participants rated ease-of-use with the System Usability Scale and the Voting Application Usability Scale. A total score was computed for each usability scale varies System Usability Scale ratings did not differ significantly among the display layouts, $F(2,22) = 1.127$, $p = 0.342$, $\eta_p^2 = 0.093$, power= 0.223 (Figure 7), nor did ratings differ on the Voting App Usability Scale, $F(2,22) = 0.716$, $p = 0.5$, $\eta_p^2 = 0.061$, power= 0.155 (Figure 8).

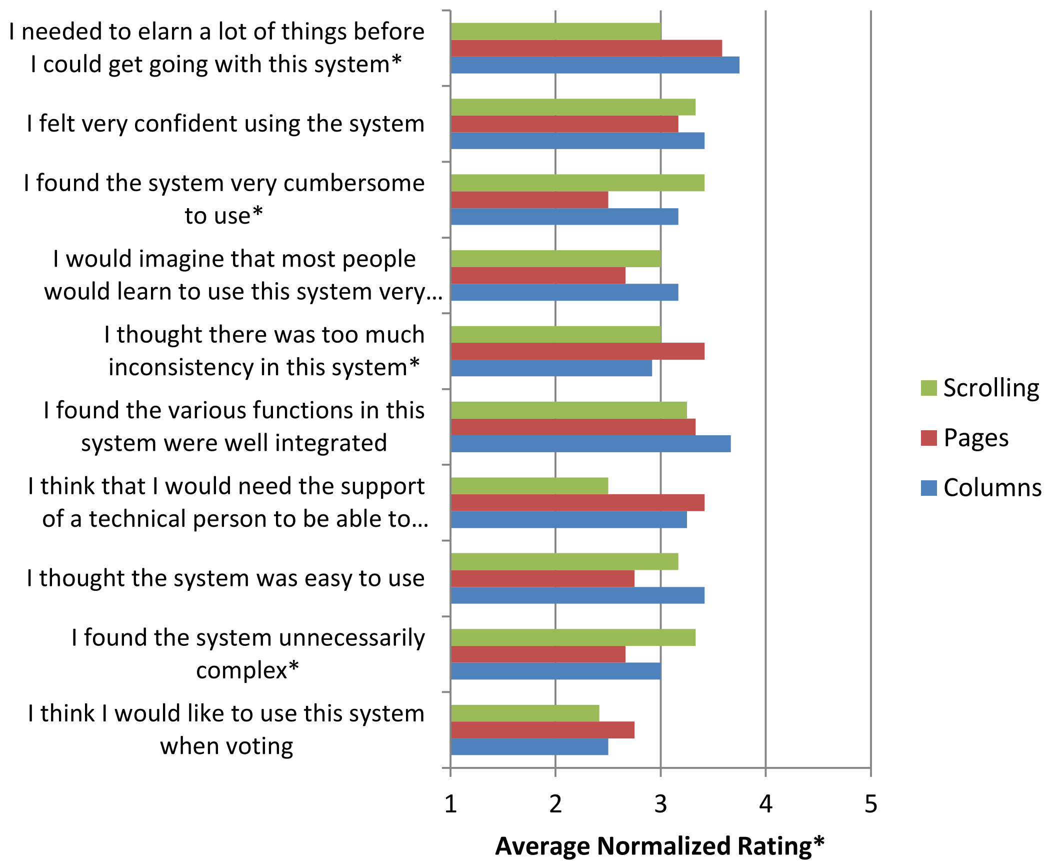

Figure 7. System Usability Scale responses for each input device. Response options ranged from 1 to 5. *For negative/undesirable statements (see starred statements), ratings were normalized so that higher scores represented greater desirability.

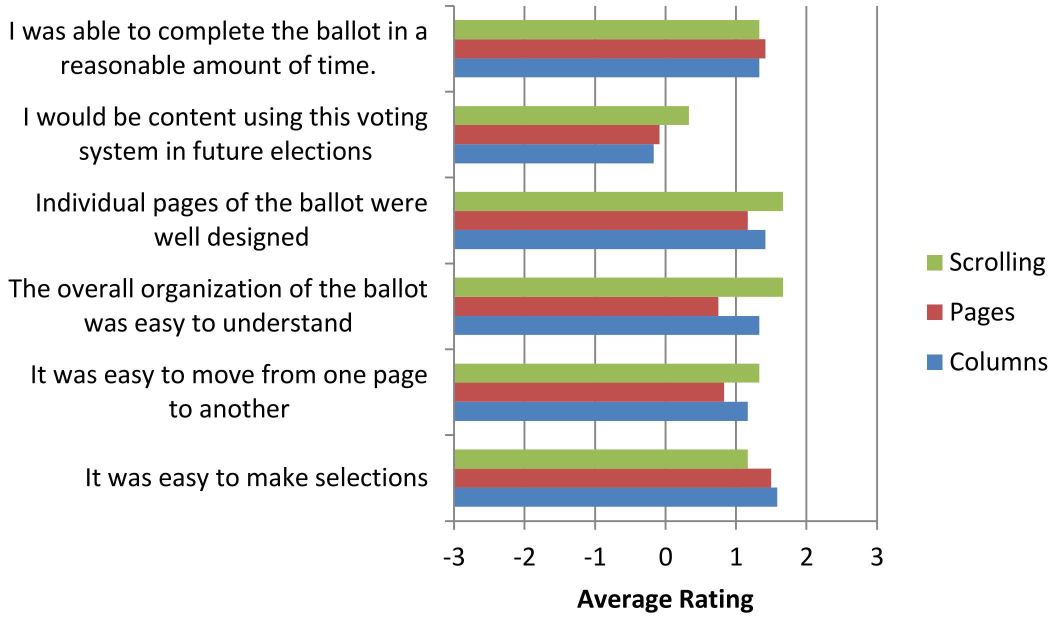

Figure 8. Voting App Usability Scale. Response options ranged from -3 to 3.

Pain

Participants indicated, using the Wong-Baker Faces Pain Rating Scale (with 0 corresponding to "no hurt" and 5 corresponding to "hurts worse"; originally published in Whaley & Wong's Nursing Care of Infants and Children), the intensity of pain they experienced while interacting with each type of display (see Figure 9). It was expected that pain ratings would be affected by the number of button presses required (H3). Repeated measures ANOVA revealed that average pain ratings do not differ significantly among the different displays, $F(2,12) = .462$, $p = .641$, $\eta_p^2 = 0.071$, power= 0.109.

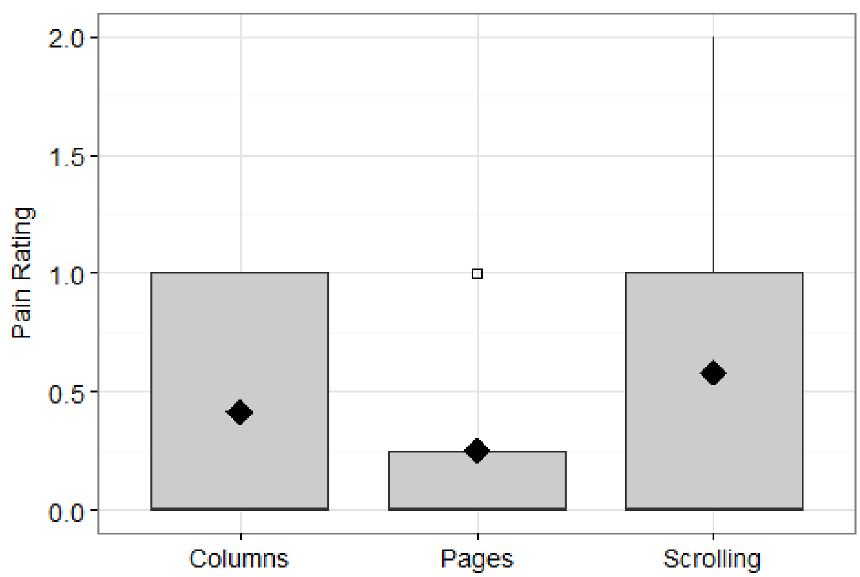

Figure 9. Pain ratings.

Objective Data

Race Time

Race time represents the amount of time it took participants to vote in the races (Figure 10). This is the sum of reaction time (time to select a candidate) and the subsequent epoch during which the participant navigated from the selected candidate to the on-screen "next" page button. It excluded the prompts before each race. Four races on each ballot were designed to contain too many candidates to display in a single column on a single page, and therefore were displayed across multi-column, multi-page, or scrolling layout. Race time was expected to be shorter on these races in the multi-column layout than in the other layouts (H1).

The results showed that race times were significantly different among the display layouts, $F(2, 22) = 7.254$, $p = 0.004$, $\eta_p^2 = 0.397$ (Figure 10). Post-hoc pairwise comparisons using the Bonferroni correction revealed that mean race time was shorter for the multi-column layout than for the multi-page layout, $p = 0.008$. No other pairwise comparisons were significant. The longer race times for the multi-page layout might have been due to confusion and an increased cognitive load; three of the 12 participants required assistance from the experimenter, who had to intervene and explain that the first page did not contain all of the candidates for the race.

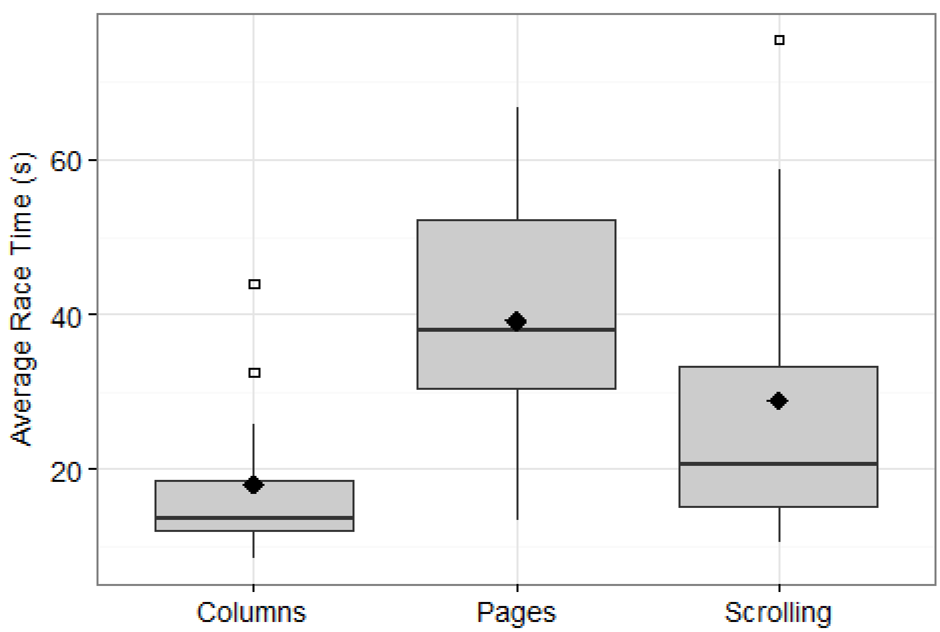

Figure 10. Participants' average race times for the four races in which the lists of candidates were too long to fit in a single column on a single page. Average time across participant is represented by 'X'.

Subtask Times

The analyses below examined two voting subtask times separately; these included moving the onscreen cursor to a desired button (*Click Time*), and finding and selecting a candidate's name (*Response Time*). Each of these subtasks reflects navigation speed, which is dependent on visual search and cursor movement.

Response Time

Response time represents the time elapsed between the onset of a race (i.e., when the list of candidates appeared) and the participant's selection of a candidate. This response time was normalized by the target candidate's location in the list. There was no significant difference among mean normalized response time for the three display layouts, F(2, 20= 0.115, p = 0.892, η_p^2= 0.011, power= 0.065 (see Figure 11).

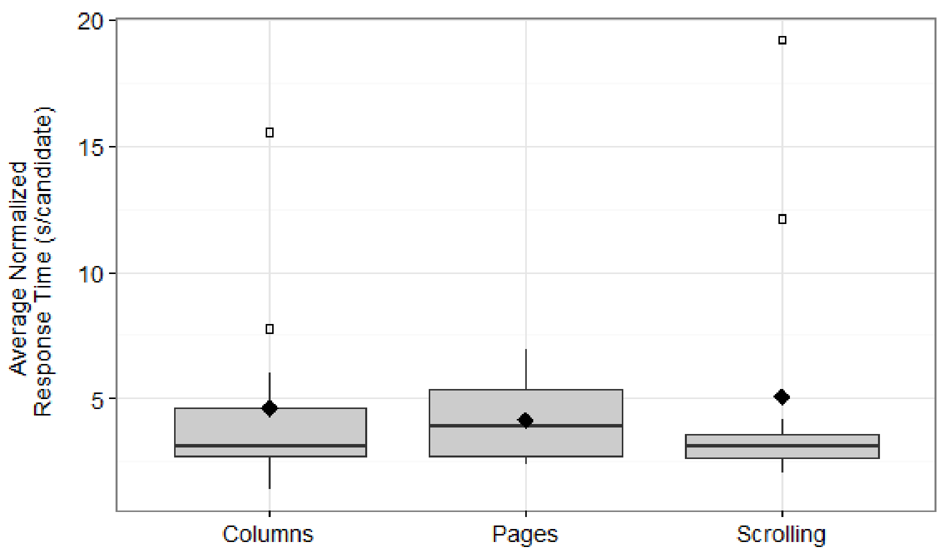

Figure 11. Participants' average response times, normalized by candidate's location in the list. Average time across participant is represented by 'X'.

Click Time

Mean click time represents the average amount of time elapsed between sections of any on-screen buttons, including navigation control buttons, and candidate names. This is an indication of the amount of time needed to move the cursor to the desired button. Repeated measures ANOVA indicated that there was no significant difference in mean click time for the three display layouts, $F(2, 22) = 2.771$, $p = 0.084$, $\eta_p^2 = 0.201$, power= 0.489 (Figure 12).

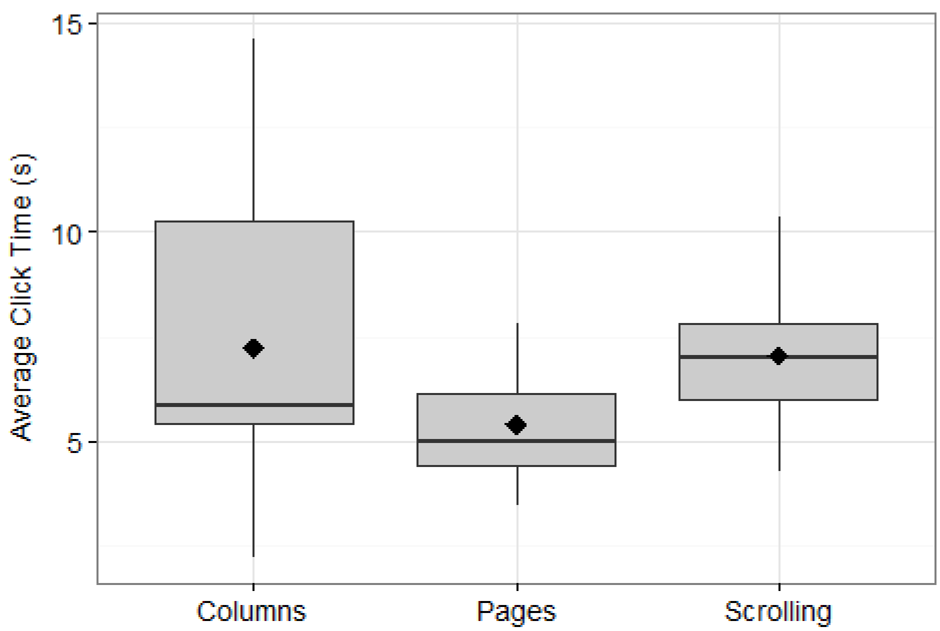

Figure 12. Average time elapsed between button selections across the entire ballot. Average time across participant is represented by 'X'.

Voting Errors

Participants were instructed to vote for a particular candidate on each race. For each race, there was only one prompted name. Participants were not prompted on ballot initiatives or retention questions; they were instructed to "vote yes or no".

Over-votes

Over-votes represent the number of times, over an entire ballot, that a participant selected more than one option in a race. It was expected that the multi-page layout would lead to more errors than the other layouts, because users might sometimes think that they have moved to another race rather than another page within the same race (H3b). Repeated measures ANOVA revealed that there was no significant difference in the average number of over-votes among display types, $F(2, 22) = 0.537$, $p = 0.592$, $\eta_p^2 = 0.047$, power= 0.127 (Figure 13). This is probably due to the large variability in occurrence of over-voting.

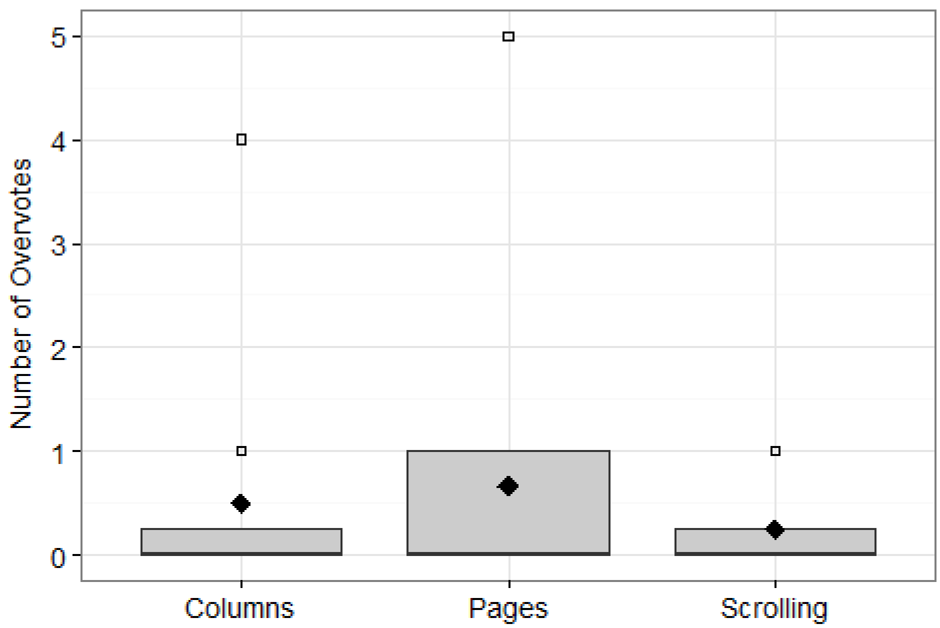

Figure 13. Participants' total number of over-vote errors per ballot.

Under-votes

Under-votes represent the number of times a participant failed to select an option (e.g., candidate) in a race. Repeated measures ANOVA indicated that there was no significant difference in the number of under-votes for the three display layouts $F(2, 22) = 2.2$, $p = 0.135$, $\eta_p^2 = 0.167$, power= 0.401 (Figure 14). However, the pattern was for multi-column layout to have the largest number of under-votes, followed by scrolling layout and then multi-page layout.

Notice that the number of under-votes is substantially higher than the number of over-votes. The experimenter observed that many participants had difficulty navigating the ballots, and this often caused under-votes. Specifically, when the cursor was on the "Next-Page" button, participants often pressed the "select" button when trying to move down to the list of candidates. This caused the ballot to advance to the next page (preceded by a pop-up under-vote warning), before the participant had a chance to select a candidate. Several participants did this repeatedly, appearing to have difficulty learning to operate the controls.

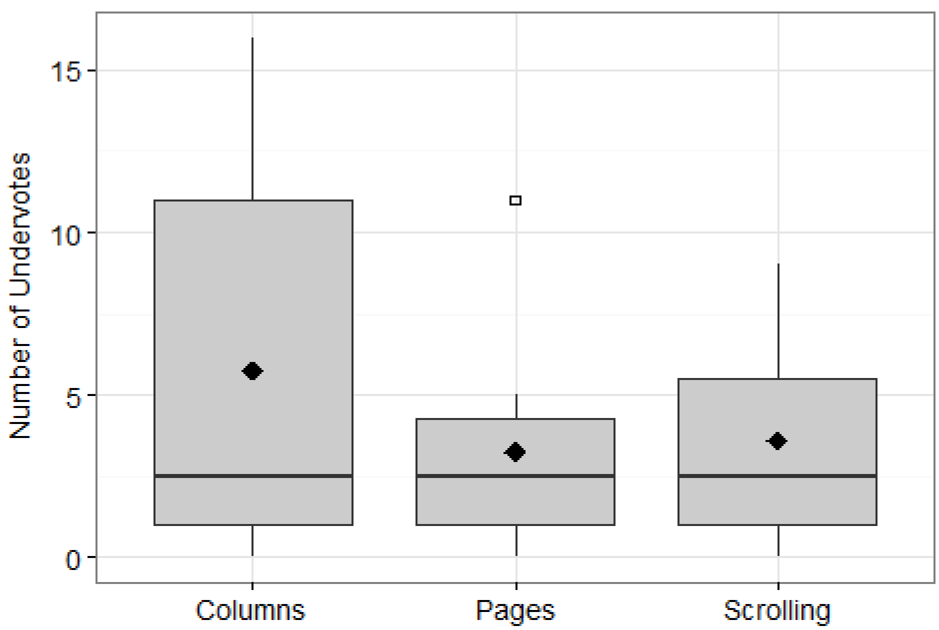

Figure 14. Participant's total number of under-vote errors per ballot.

Percent Correct

Percent correct indicated the percentage of correctly chosen candidates. Participants were prompted to vote for one candidate on each of 12 races. Repeated measures ANOVA indicated that there was no significant difference in percentage of correctly chosen candidates among the three display layouts, $F(2, 22) = 0.299$, $p = 0.745$, $\eta_p^2 = 0.026$, power= 0.091 (Figure 15). This indicates that participants were able to accurately identify the target candidate that they were primed to vote for.

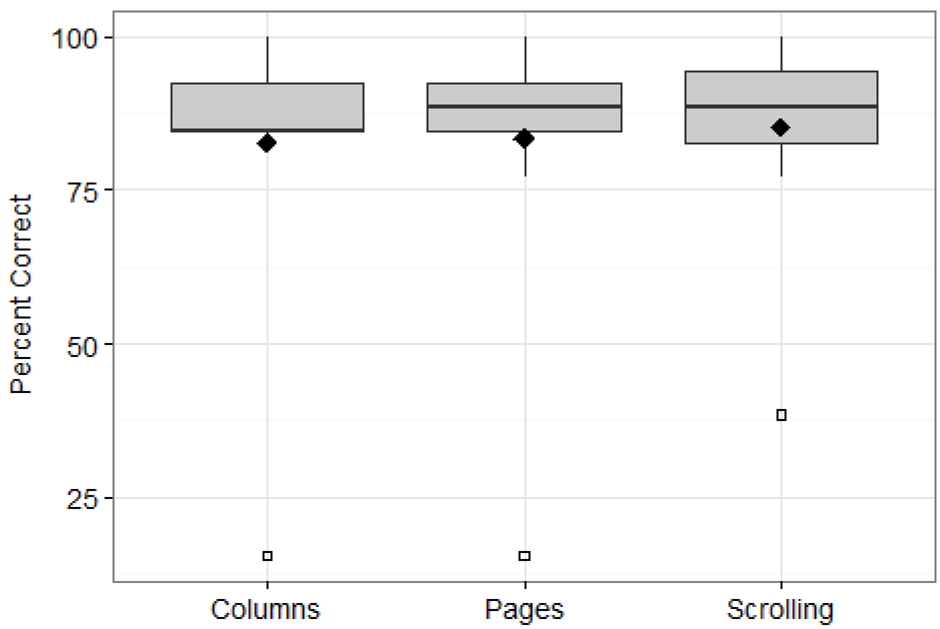

Figure 15. Percent correct was calculated as a percentage of time that participants voted for the prompted candidate The three outliers are from one participant who do not comply with the prompts.

Eye Tracking Results

Eye tracking data was collected for high-content races on which candidates were displayed differently on the three ballots. All values below represent summed gaze times over these four races.

Controls Gaze Time

Control gaze time refers to the amount of time participants spent looking at the on-screen navigation control buttons at the top of the ballot. Repeated measures ANOVA revealed that there was a significant difference in the time spend looking at the on-screen navigation control buttons among the three layouts, $F(2,14) = 3.773$, $p = .049$, $\eta_p^2 = 0.589$. However, Mauchley's test for violation of sphericity was significant, $X^2(2) = 8.409$, $p = .015$. With the Greeenhouse-Geisser correction for sphericity violation, the difference among means was not significant, $F(1.140, 7.983) = 3.773$, $p = 0.085$, $\eta_p^2 = 0.421$. Although none of the post-hoc pairwise comparisons were statistically significant, the average gaze time on the controls was highest for the multi-page layout. This is consistent with the hypothesis that participants would have more difficulty navigating the ballot in the multi-page layout (H3a).

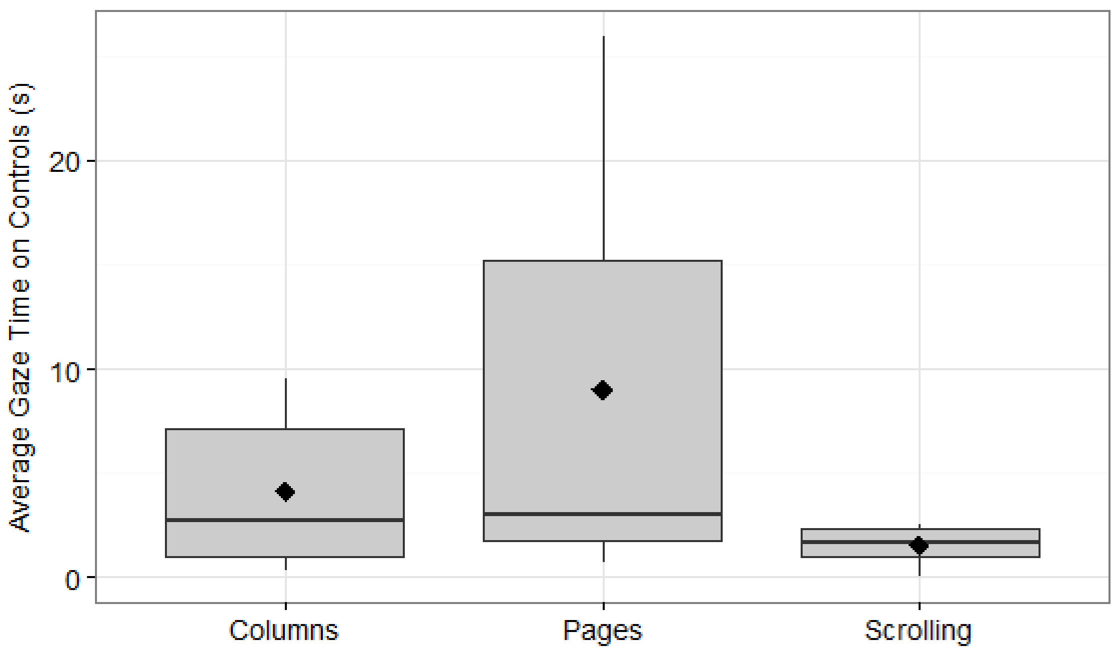

Figure 16. Mean time per page spent looking at the on-screen navigation controls.

Candidate Gaze Time

Candidate gaze time refers to the amount of time participants spent looking at all candidate options on the ballot. An average was computed across the relevant races within each ballot. Repeated measures ANOVA revealed that there was a significant difference in time spent looking at the candidate options among the three layouts, $F(2, 14) = 12.249$, $p = 0.001$, $\eta_p^2 = 0.636$ (Figure 17). Post-hoc analysis using a Bonferroni correction revealed that participants looked at candidates on the multi-page layout for significantly longer time than either the multi-column layout ($p = 0.005$) or the scrolling layout ($p = 0.016$). The difference between the multi-column and scrolling layouts was not significant. This is consistent with the hypothesis that participants would have more difficulty navigating the ballot in the multi-page layout (H3a).

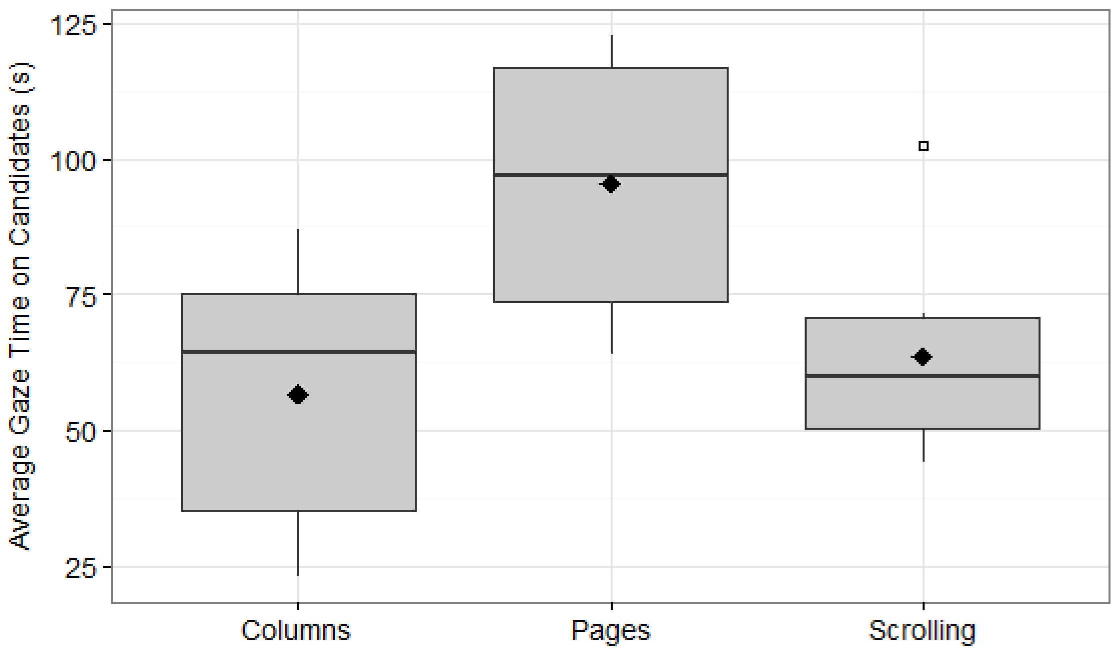

Figure 17. Participants' average times per race spent looking at the list of candidates.

Visual Search Time

Participants were expected to take longer to visually search for the target candidate in the multi-page layout than the other two layouts (Hypothesis H2). Visual search time refers to the amount of time participants spent looking at the candidate options, excluding the target candidate. This provides an estimate of the amount of time it took participants to find the target candidate.

Repeated measures ANOVA revealed there was a significant difference in the visual search time among layouts, $F(2, 14) = 14.939$, $p < 0.001$, $\eta_p^2 = 0.681$ (Figure 18). Post-hoc analyses using a Bonferroni correction revealed that visual seaerch time was significantly longer on the multi-page layout than either the multi-column layout ($\rho = 0.003$) or the scrolling layout ($p = 0.015$). The scrolling and multi-column layouts were not significantly different. Participants may have searched the candidate list in the multi-page layout longer as they searched for the prompted candidate's name, not realizing that the name might have been on another page.

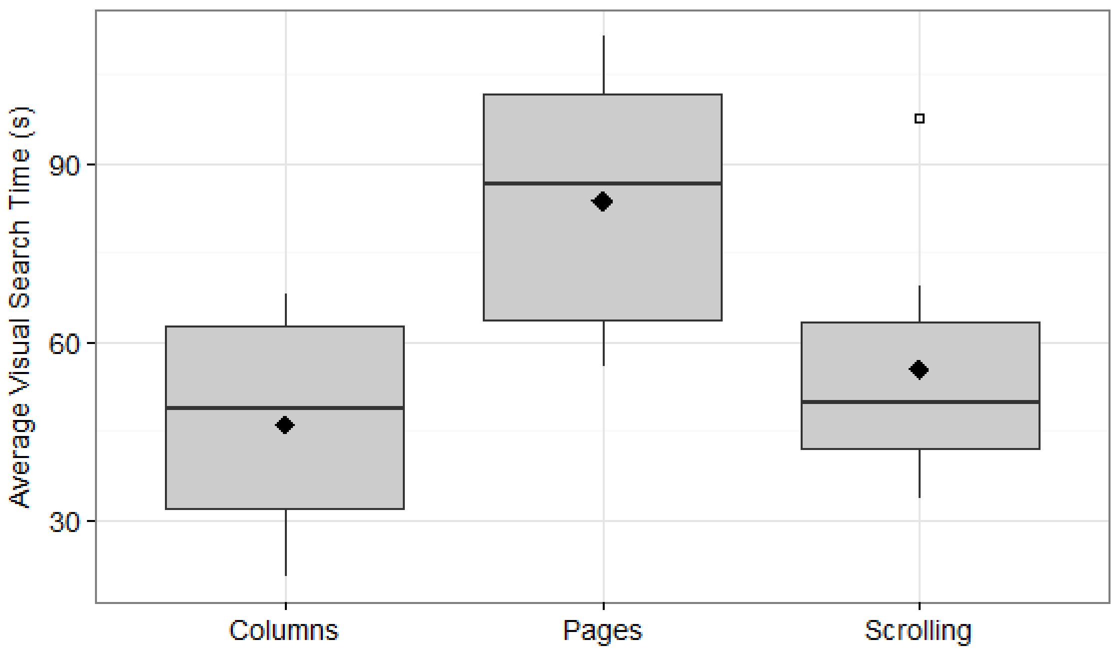

Figure 18. Mean time per race spent looking at non-target candidates while searching for the target candidate.

Target Candidate Gaze Time

Target candidate gaze time refers to the amount of time participants spent looking at the target candidate. It was expected that once the target candidate is found that there would be no difference in how that name is displayed across the various layouts. Repeated measures ANOVA revealed that there was no significant difference in the amount of time spent looking at the target candidate among the ballot types, $F(2, 14) = 0.514$, $p = 0.609$, $\eta_p^2=0.068$, power= 0.118 (Figure 19). Once participants were able to find their target candidate, there was no difference in how long it took them to recognize that they had found the right candidate.

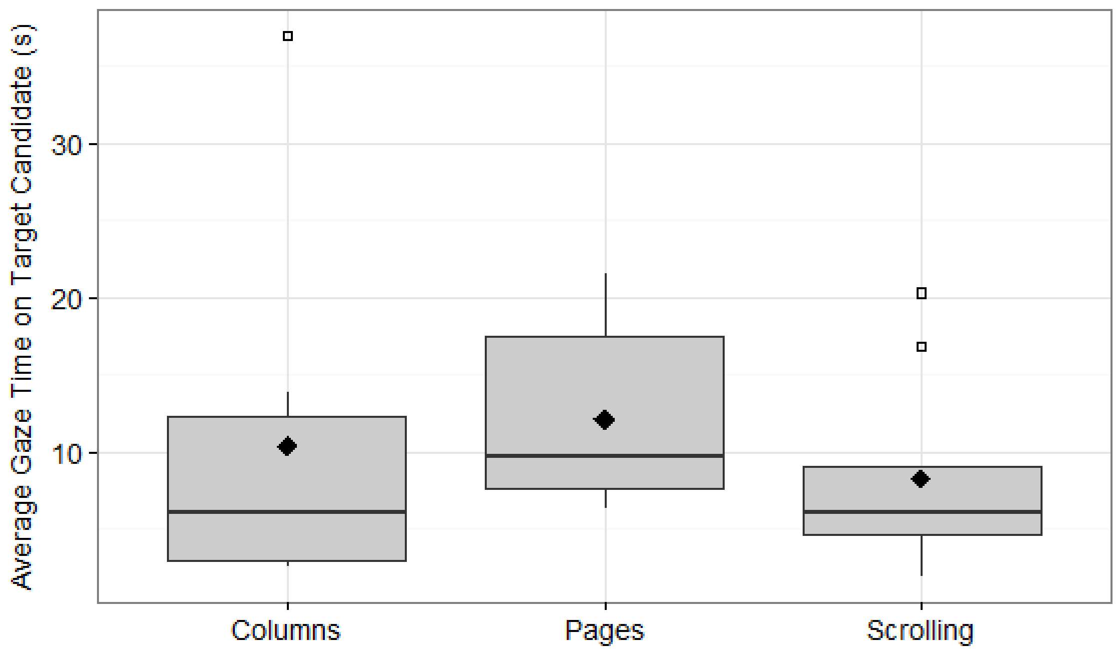

Figure 19. Participant's average time per race spent looking at the target candidate.

Discussion

The purpose of this research study was to determine which of the three different display layouts would be best suited for older adults with arthritis in their wrists, hands, and/or fingers using a 3-button controller. The three display layouts that were evaluated were as follows: multi-page, multi-column, and scrolling, and these layouts were leverage for candidate races that were exceptionally long and unable to fit into a single column on a single page.

The multi-page layout was considered the worst option, in that participants took a significantly longer time to complete the ballot (race time), visually search for candidates, and it received the lowest ranking (although not statistically significant). The multi-page layout may have caused higher cognitive load, because participants had to use their working memory to recall the targeted candidate across various pages, and had to conceptually understand that a race could span multiple pages. For all of the variables considered there were no significant differences between the multi-column and scrolling layouts.

Participants tended to spend more time looking at the on-screen navigation control area for the multi-page layout than the multi-column or scrolling layout. In addition participants took a significantly longer amount of time to find the targeted candidate for the multi-page layout than the multi-column and scrolling layouts. This is an indication, that the visual search task was more complex for the multi-page layout, and is possibly why 7 out of the 12 participants ranked multi-page layout as the worst option. Little differences in gaze times were observed between the multi-columns and scrolling layouts.

Overall, multi-page layout was the least desirable configuration, while there were little difference between multi-column and scrolling layouts. The primary benefit of multi-column layout, as observed by participants, is that all of the candidates are visible at a glance. However, there are certain conditions that may exist that prohibit using multi-column layout effectively. For example, if the list of candidates is too long in order to maintain readability in a multi-column layout. In this instance, it is advisable to instead display the candidate list in a scrolling layout. However, it is not recommended to utilize the multi-column and scrolling layout simultaneously, as this would increase the cognitive load. If scrolling layout is used, it is recommended to have an indicator to notify the user where they are in the list, and also to split the bottom/top button/candidate name in half, to indicate that more information is available below or above the current location. These visual cues should facilitate making scrolling easier to use. If these two design features are included, then according to the present study results scrolling layout would be as usable as multi-column layout. Therefore the final recommendation is to use multi-column layout where appropriate, but if the list of candidates or referendum is too long to then make use of scrolling as opposed to placing the information across multi-page layout.

One of the limitations of this study was sample size, in that data from only twelve participants were collected. The statistical power for all of the non-significant analyses was below 0.5, compared to the typically desired power of 0.8. Also it is somewhat difficult to generalize the results to other populations, since we only considered older adults with arthritis. In addition, since we did not measure working memory it is difficult to ascertain whether this would have influenced their preferences, and identify any potential outliers. Therefore, in a sense the results from this study is a good control, future research studies could evaluate cognitive functional capability and user interface design. Lastly, these results are relevant to individuals using a tactile user input device, as opposed to touch screen or mouse interface. This may affect the intuitiveness of using the system and potentially result in a different perspective between the multi-column and scrolling layouts. However, this too is a potential future research question that should be addressed.

In conclusion, it is recommended to utilize a multi-column layout when appropriate, followed by the scrolling layout if the lists of candidates are too long. However, it is recommended to refrain from splitting candidate lists for a single contest across multiple pages, as this increases the complexity of navigating the ballot.

Acknowledgments

The authors would like to thank Dr. Cara Bailey Fausset and Ms. Hannah Jahant for their contributions in reviewing and editing this manuscript.

References

Andrevey, V., Martynov, A. (2000) Effects of splitting text into multiple columns. *Dept. of Computer Science Class Project,* Univ. of Maryland, College Park.

Bouma, H. (1980) Visual reading processes and the quality of text displays. Vigliani, E. (eds.) *Ergonomic Aspects of Visual Display Terminals*, Taylor & Francis, London.

Chaparro, B., Baker, J.R., Shaikh, A.D., Hull, S., Brady, L. (2004) Reading online text: A comparison four whitspace layouts. *Usability News* 6.2.

De Brujin, D., De Mul, S., Van Oostendorp, H. (1992) The influence of screen size and text layout on the study of text. *Behavior and Information Technology* 11:71 – 78.

Duchnicky, R.L., Kolers, P.A. (1983) Readability of text scrolled on visual display terminals as a function of window size. *Human Factors* 25:683 – 692.

Dyson, M.C., Kipping, G.J. (1997) The legibility of screen formats: Are three columns better than one? *Computers and Graphics* 21(6):703 – 712.

Dyson, M.C., Kipping, G.J. (1998) The effects of line length and method of movement on patterns of reading from screen. *Visible Language* 32:150 – 181.

Dyson, M.C., Haselgrove, M. (2001) The influence of reading speed and line length on the effectiveness of reading from screen. *Int. J. Human-Computer Studies,* 54:585 – 612.

Hall, T., & Snoot, L. W., (2013). Voting Machines: The Question of Equal Protection. In M. J. Streb (Ed.), Law and Election Politics: The Rules of the Game, 2nd ed. 4, 81 – 87, New Your, NY: *Lynn Renier Publishers*.

Hambrick, D. Z., & Engle, R. W. (2002). Effects of domain knowledge, working memory capacity, and age on cognitive performance: An investigation of the knowledge-is-power hypothesis. [Article]. *Cognitive Psychology, 44*(4), 339-387.

Harada, E. T., Mori, K., & Taniue, N. (2010). Cognitive aging and the usability of IT-based equipment: Learning is the key. [Article]. *Japanese Psychological Research, 52*(3), 227-243.

Harley, L.R., Mosley, S., Jones, A., Baranak, A., Kline, K., and Fain, B. (2013). A web based voting application software design. *ITIF Working Paper*.

Lam, K., Lam Y., Liu, J., Shin, U.G. (2000) Reading comprehension and rate: one column versus three columns. *Dept. of Computer Science Class Project*, Univ. of Maryland, College Park.

Lawrence, R. C., Helmick, C. G., Arnett, F. C., Deyo, R. A., Felson, D. T., Giannini, E. H., . . . Wolfe, F. (1998). Estimates of the prevalence of arthritis and selected musculoskeletal disorders in the United States. *Arthritis & Rheumatism, 41*(5), 778-799.

Morbidity and Mortality Weekly Report (2013). Prevalance of doctor-diagnosed arthritis and arthritis-attributable activity limitation – United States, 2010 – 2012, *Centers for Disease Control and Prevention MMWR* 62(44).

Sauro, J. (2011). Measuring usability with the system usability scale Retrieved February 25, 2013, from http://www.measuringusability.com/sus.php