Copyright
Danae Holmes
2014
RICE UNIVERSITY

Checking Your Answers: An Investigation of Alternative Review Screen
Design for Electronic Voting Systems

by

Danae V. Holmes

A THESIS SUBMITTED
IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE

Doctor of Philosophy

APPROVED, THESIS COMMITTEE

Philip T. Kortum, Chair
Assistant Professor, Psychology

Michael D. Byrne
Professor, Psychology and Computer Science

David M. Lane
Associate Professor, Psychology, Statistics, and Management

Dan S. Wallach
Professor, Computer Science and Electrical & Computer Engineering

HOUSTON, TEXAS
2014
ABSTRACT

Checking Your Answers: An Investigation of Alternative Review Screen Design for Electronic Voting Systems

by

Danae Holmes

Verifying a ballot for correctness in an election is a critical task considering the large, negative repercussions of an incorrect ballot. Studies have shown weaknesses in the ballot review process in electronic voting systems, allowing up to 30% of the ballot to be changed without being noticed by 68% of voters (Everett, 2007). There is also a noted lack of research on the effect of navigation style on electronic voting system usability and review screen performance. In response to these issues, this study evaluated the usability and viability of alternative ballot verification and navigation methods in an electronic voting medium, specifically direct recording electronic (DRE) voting systems. Currently, most DREs employ an end-of-ballot review where all selections are confirmed at once at the end of the ballot, which has been shown to be ineffective. Several studies (Holmes & Kortum, 2013; Selker 2007) have also indicated that in-line confirmation, confirming each selection immediately after making it, and a combination of the two confirmation methods (Ghandi et al., 2005; Cohen et al., 1996) may prove to be a suitable alternatives.

The current study tested these methods of verification in terms of performance and usability to determine whether they are viable methods of verification as well as to provide a benchmark for review screen performance in a DRE. The method of navigation through the ballot, the ability to move backwards through the ballot or not after selecting a candidate, was
also tested for its impact on usability and performance. The verification methods were evaluated on three metrics of usability as defined by ISO 9241 part 11: efficiency (time to complete a ballot), effectiveness (errors), and satisfaction (subjective usability). Participants cast their ballot in a mock national election using a custom DRE interface. Results indicate that in-line and dual confirmation methods may be viable alternatives for DRE review screens. In-line and dual confirmation perform similarly to end-of-ballot confirmation in terms of effectiveness, but differed in other usability and performance aspects, though not necessarily in a negative way. The most efficient method is end-of-ballot review, and dual confirmation produced the longest time spent on the review screen. End-of-ballot confirmation produced the highest satisfaction ratings, though survey results indicated that dual confirmation may be the most appropriate method in terms of voting. Based on the results from this study, further studies should be conducted to determine which confirmation method performs best as an error prevention tool.
DEDICATION

This dissertation is dedicated to my father and mother, Darrell and Sharon Holmes, and Robert De Jesus, for their incredible support during my journey through graduate school. I have come this far only because of your unrelenting love and encouragement.
AKNOWLEDGMENTS

I would first and foremost like to thank my advisor, Dr. Philip Kortum, for his patient guidance throughout my years at Rice University. He has successfully gotten me through the program as well as trained me to be an effective human factors practitioner and researcher. I am very thankful for all the time he invested in guiding me through the doctoral program.

I would also like to thank the members of my dissertation committee: Dr. Mike Byrne, Dr. David Lane, and Dr. Dan Wallach. Their feedback significantly improved the quality of research in this dissertation.

I want to thank all of my colleagues in the Department for their help, input, and friendship during my time at Rice. I would like to thank Claudia Acemyan, Heather Dial, Chad Tossell, Lanita Martin, and Jennifer Gucwa, for their help in getting me through the doctoral program, especially through the first few years.

I want to thank God, who has guided me throughout my life and has given me wisdom and strength during my graduate education. I am especially thankful for the support my family and friends have given me, for without them my success would not have been possible. A special thanks to Mara Jackson, Anna Hofmanova, and Danielle Brager, for keeping me grounded and for their encouragement during graduate school.
# TABLE OF CONTENTS

DEDICATION......................................................................................................................... iv

AKNOWLEDGMENTS .............................................................................................................. v

LIST OF FIGURES .................................................................................................................. x

LIST OF TABLES ................................................................................................................... xii

Introduction............................................................................................................................ 1

DRE Voting Systems ............................................................................................................. 1

Election Integrity .................................................................................................................. 2

Lack of Research on DRE Review Screens ......................................................................... 4

Ineffective DRE Review Screen .......................................................................................... 4

Alternative Approaches ......................................................................................................... 9

Defining Current and Alternative Review Methods .............................................................. 10

End-of-Ballot ......................................................................................................................... 10

In-Line ................................................................................................................................ 10

Dual ...................................................................................................................................... 10

Strengths and Weaknesses .................................................................................................. 11

Efficiency ............................................................................................................................... 11

Cognitive Demand ................................................................................................................. 13

The “Fleeing Voter” Phenomenon ......................................................................................... 18
Previous Implementation ................................................................. 19
Redundancy................................................................................... 22
Effect of Navigation on Review .................................................... 23
Study ......................................................................................... 25
Method ....................................................................................... 26
Participants ................................................................................. 26
Design ......................................................................................... 27
Materials ..................................................................................... 28
Apparatus .................................................................................... 28
Other Materials .......................................................................... 34
Procedure .................................................................................... 35
Results ....................................................................................... 35
Usability Results ......................................................................... 36
Efficiency ..................................................................................... 36
Effectiveness .............................................................................. 37
Satisfaction ................................................................................ 39
Review Screen Time ................................................................... 41
Vote Order .................................................................................. 44
Verification Method Survey ....................................................... 48
Discussion .................................................................................. 48
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Usability</td>
<td>48</td>
</tr>
<tr>
<td>Efficiency</td>
<td>49</td>
</tr>
<tr>
<td>Effectiveness</td>
<td>50</td>
</tr>
<tr>
<td>Satisfaction</td>
<td>50</td>
</tr>
<tr>
<td>Performance and Design</td>
<td>51</td>
</tr>
<tr>
<td>Qualitative Results</td>
<td>53</td>
</tr>
<tr>
<td>Limitations</td>
<td>54</td>
</tr>
<tr>
<td>Future Directions</td>
<td>56</td>
</tr>
<tr>
<td>Conclusion</td>
<td>56</td>
</tr>
<tr>
<td>References</td>
<td>58</td>
</tr>
<tr>
<td>Appendix A</td>
<td>65</td>
</tr>
<tr>
<td>VoteBox Lite Design</td>
<td>65</td>
</tr>
<tr>
<td>VoteBox Lite Screenshots</td>
<td>71</td>
</tr>
<tr>
<td>Appendix B</td>
<td>75</td>
</tr>
<tr>
<td>Background and General Voting Survey</td>
<td>75</td>
</tr>
<tr>
<td>Confirmation Method Survey</td>
<td>80</td>
</tr>
<tr>
<td>Ballot</td>
<td>83</td>
</tr>
<tr>
<td>Republican Slate</td>
<td>85</td>
</tr>
<tr>
<td>Democratic Slate</td>
<td>87</td>
</tr>
<tr>
<td>Appendix C</td>
<td>89</td>
</tr>
</tbody>
</table>
LIST OF FIGURES

Figure 1. Image of the Hart InterCivic eSlate DRE (left) and Diebold AccuVote-TS DRE (right). .............................................................................................................................................................................................................................................................................................................. 2

Figure 2. Depiction of contest on VoteBox from Everett (2007). ................................................................................................. 5

Figure 3. Depiction of review screen displayed on VoteBox from Everett (2007). ......................................................... 6

Figure 4. Updated VoteBox review screen from Campbell and Byrne (2009). ............................................................ 8

Figure 5. Diagram depicting the concept of in-line confirmation. .............................................................................................. 10

Figure 6. An example of dual confirmation in a DRE. ...................................................................................................................... 11

Figure 7. Screenshot of the end-of-ballot review screen in VoteBox Lite. .......................................................... 29

Figure 8. Screenshot of VoteBox Lite in-line confirmation screen with a contest selection..... 30

Figure 9. Screenshot of VoteBox Lite in-line confirmation screen without contest selection. .... 31

Figure 10. VoteBox Lite in-line confirmation ballot submission screen............................................................................. 31

Figure 11. Screenshot of a contest with back navigation button highlighted. .......................................................... 32

Figure 12. Screenshot of a contest without back navigation button. ...................................................................................... 33

Figure 13. Screenshot of VoteBox Lite PIN screen.......................................................................................................................... 34

Figure 14. Screenshot of VoteBox Lite voting instructions screen. ......................................................................................... 34

Figure 15. Mean ballot completion time in seconds by verification type................................................................. 36

Figure 16. Comparison of efficiency metric between different voting systems. .......................................................... 37

Figure 17. Mean error rate for each verification type. Error bars represent standard error. ............. 38

Figure 18. Mean error rate comparison of VoteBox Lite with other voting methods. ......................... 38

Figure 19. Mean SUS scores for each verification method. ................................................................................................. 40

Figure 20. Mean SUS scores for VoteBox Lite and other voting methods. ......................................................... 40

Figure 21. Mean time spent on the review screen for each verification method. ............................................. 42
Figure 22. Review times for in-line and end-of-ballot verification methods compared to their respective counterparts for dual verification........................................................... 43

Figure 23. Mean review time for each race with in-line verification. ........................................ 43

Figure 24. Mean review time for each race with the in-line confirmation portion of dual verification. ......................................................................................................................... 44

Figure 25. Interaction between verification type and vote order on efficiency. ......................... 45

Figure 26. Interaction between verification and vote order on satisfaction................................. 46

Figure 27. Time to complete a ballot for first, second, and third trial ........................................ 47
LIST OF TABLES

Table 1. Distribution of Participant Ethnicities ................................................................. 27
Table 2. Frequency of Errors by Ballot .................................................................................. 39
Table 3. Efficiency Interaction Means .................................................................................. 45
Table 4. Satisfaction Interaction Means ................................................................................ 46
Table 5. Verification Strengths and Weaknesses.................................................................. 54
Introduction

Verifying information is a common task performed in a variety of situations such as checking test answers, reviewing receipts, and even deleting files on a computer. Verification is vital for tasks like transferring funds online or trading stock, where errors could lead to very serious consequences. Voting is an area where verification of accuracy is critical. This is especially true with electronic voting systems, where there is a potential security threat to ballot integrity (Bannet, Price, Rudys, Singer, & Wallach, 2004).

Electronic voting systems typically offer a first line of defense against ballot errors and malicious ballot tampering in the form a review screen. These screens allow users to review and confirm or change their choices before submitting the ballot. Several studies have shown that review screens in direct recording electronic (DRE) voting systems are not optimal in helping users detect changes or errors in their ballot (Everett, 2007; Campbell & Byrne, 2009; Acemyan, Kortum, & Payne, 2013). There is also a lack of research on the impact of global navigation types on DRE review screen usability. The goal of this study was to explore different methods of confirmation as well as the interaction between ballot navigation and DRE review to find viable alternatives to the current method of review. The implications of this study are wide-reaching in that they could not only improve the efficacy of electronic voting system review screens, but verification in similar systems, like online banking transactions, as well.

DRE Voting Systems

First, it is essential to gain a firm understanding of DRE voting systems. DREs are the newest form of voting system, with traditional methods being paper ballots, lever machines, and punch cards. Though electronic voting machines seem like a recent development, the first patents for these machines actually emerged around the 19th century (Jones, 2003). The first electronic
voting machines became commercially available in 1974 (Jones, 2003). Since that time, DREs have evolved into the machines we currently use today as seen in Figure 1.

![Image of voting machines](image)

**Figure 1.** Image of the Hart InterCivic eSlate DRE (left) and Diebold AccuVote-TS DRE (right).

DREs present a digital version of a ballot and allow voters to make their selections via various input methods such as touchscreens and physical navigational buttons. These systems also provide a self-check mechanism in the form of a review screen. This screen allows voters to review their selections for errors to ensure the accuracy of their ballot before submission. Other than that, DREs vary in terms of design, including the presentation of the ballot, ballot navigation and selection mechanisms, screen size, and review screen (Everett, 2007).

**Election Integrity**

The 2000 election in Florida revealed the serious impact that unintentional ballot errors could have on the outcome of an election. In Palm Beach County, Florida, over 2,000 voters inadvertently voted for Pat Buchanan and not Al Gore as intended due to the design of the butterfly ballots used in the election (Wand, et al., 2001). This pushed the outcome of the election in favor of George W. Bush (Wand, et al., 2001). Mebane (2004) also found that due to
unintended overvote errors (selecting more candidates than allowed) approximately 50,000 votes for George W. Bush or Al Gore were forfeited. Mebane (2004) suggested this issue could have been avoided if there were technology available that could make voters aware of the errors. Malicious attacks on a small scale can also have a significant effect on election integrity. Di Franco, Petro, Shear, and Vladimirov (2004) conducted a study showing how small manipulations in elections, such as vote flipping, could change the outcome of an election. In their study, it was shown that a close election, like the 2000 Presidential election, could have a different outcome with just one vote per voting machine being changed. Ballot verification becomes very important in situations such as this, as the impact of even small ballot changes or errors could be substantial in an election.

The Help America Vote Act (HAVA) of 2002 was enacted as a response to these election issues. The HAVA helped provide funding for states to purchase reliable, accessible, and more usable voting systems so that the problems faced in the 2000 election would not happen again (United States Government, 47th Congress, 2002). This led to many states purchasing DRE voting systems to replace older voting technologies (Everett, 2007). These systems generally permit voters to verify their ballots for accuracy and help prevent error by allowing voters to check their selections via a review screen, similar to what Mebane (2004) had suggested.

Though DREs are not the most prevalent method of voting, a considerable number of jurisdictions use them in elections. The US Election Assistance Commission estimated that 18 states in the US used DREs with voter-verified paper audit trail or VVPAT (provides a paper record of the voter’s ballot for verification) in the 2012 election (U.S. Election Assistance Commission, 2013). This figure did not include states that used DREs without VVPAT, so the usage of DREs may have been more widespread. The current usage of DREs coupled with
potential election integrity issues, both intentional and unintentional, stresses the significance of ineffective DRE ballot review.

**Lack of Research on DRE Review Screens**

Research has been conducted on the usability of DREs and other traditional voting methods by Greene, Byrne, and Everett, 2006, but little research has been done on the actual effectiveness and design of review screens (Everett, 2007). According to Fischer and Coleman (2005), DREs have not undergone the extensive scientific analyses and review as expected for voting systems. Alvarez (2002) states that a common criticism of electronic voting systems is that review screens are poorly designed. Alvarez (2002) also comments on the many considerations that must be made in the design of review screens such as highlighting undervotes, which are races in which voters intend to vote but do not, and listing candidates’ political party affiliations. These decisions should be made according to scientific research on proper review screen design, but cannot due to the lack of literature and empirical research. It is noted by Bell, et al. (2013) that the current Voluntary Voting System Guidelines version 1.1 (VVSG) needs additional information pertaining to the design of review screens. This evidence collectively shows that there is a lack of literature on how to properly design electronic voting system review screens.

**Ineffective DRE Review Screen**

Though there is lack of research on review screen design, research evaluating the effectiveness of DRE review screens exists. These studies revealed a large weakness in current DRE review screen design and stressed the serious consequences of poor design.

Everett (2007) conducted a study showing exactly how important the DRE review screen is in the accuracy of recording voter intent. This was done by reviewing the usage of DRE
review screens and how users performed when their votes were changed. The study asked the question of whether the added security of the review screen for DREs is effective, since their purpose is to give users the opportunity to review their votes for accuracy before submitting their ballots. This study included two experiments that both tested the effectiveness of the review screen by looking at whether or not voters noticed changes made to their selections on the review screen. Everett (2007) used a prototype DRE voting system called VoteBox. VoteBox was developed by Sandler, Derr, and Wallach (2008) and tested for usability against other voting methods in studies by (Everett, et al., 2008). VoteBox displays one race at a time on the screen and has a review screen at the end of the ballot. An example of how races are presented and the review screen of VoteBox can be seen in Figure 2 and Figure 3 respectively.

Figure 2. Depiction of contest on VoteBox from Everett (2007).
One experiment in Everett’s (2007) study focused on whether or not subjects would notice either an addition or removal of contests on the review screen. The number of races either added or removed varied from 2 or 8 on a 27 item ballot. The time spent on the review screen and whether or not subjects noticed the changes on the review screen, based on self-report, were recorded.

About 68% of subjects did not notice changes in their ballot on the review screen. The number of races added or removed did not affect the chance of subjects noticing the change. Their data showed that up to 30% of the ballot could be altered without being noticed by the majority of users. Everett (2007) suggested that this effect would likely be seen with even higher numbers of changed contests. The amount of time a subject spent on the review screen was a good predictor of whether or not they would detect changes. The longer a subject spent looking at the review screen, the more likely they were to detect errors.
The second experiment was similar to the first in that it focused on whether or not subjects used the review screen and if they noticed changes to their ballots on the review screen. Instead of adding or removing contests, the main manipulation was changing contest selections, often called vote flipping. Votes could be changed to another candidate or to “none,” which would represent an undervote. There were either 1, 2, or 8 changes to contest selections. The locations of the changes were at the top or bottom of the ballot with varying positions of top or bottom of the ballot. This allowed for an examination of whether the position of the error on the ballot affected change notice.

The results of this experiment were similar to the first experiment in that the majority of subjects did not notice changes in their ballot on the review screen. About 63% of users did not notice changes on the review screen. The number and location of changes did not have an effect on whether subjects noticed the changes. This means that even the more important races, such as for US Senator, could be changed with a small chance of detection. Similar to the first experiment, there was a correlation between the amount of time spent on the review screen and the detection of errors.

Currently, most DRE systems have review screens similar to that of VoteBox’s that are displayed at the end of the ballot (Everett, 2007). Everett’s (2007) study showed that this style of review screen may not be an effective error prevention tool. This could lead to large security issues in that voters’ ballots could be changed with the majority never even noticing. It also suggests that users may not be able to detect their own errors on a ballot.

Campbell and Byrne (2009) extended Everett’s (2007) research by making changes to the VoteBox review screen and subjects’ instructions to ascertain if this increased users’ awareness of changes. Campbell and Byrne (2009) identified potential improvements that could be made to
review screens that might increase their performance. These improvements included increasing the salience of undervotes (contests with no selections), introducing political party information on the review screen, and adding visible instructions in VoteBox stressing the importance of review screen verification (see Figure 4). In this study, subjects were made aware of the importance of review screen verification. The salience of undervotes was increased via highlighting. The main manipulations and materials in this study were the same as Everett’s (2007). The number of flips were either 1, 2, or 8 and the position of the flips varied from either the top or bottom of the ballot.

Figure 4. Updated VoteBox review screen from Campbell and Byrne (2009).

This study indicated that the review screen changes. About 50% of subjects did not notice changes in their ballot compared to the 63% found in Everett (2007). Most of the effects seen in Everett’s (2007) study were replicated in the results as well. These included the effect of time
spent on the review screen on change notice and the fact that the number and position of changes had no effect on change noticing.

The results of this study have large implications for the importance of review screen design in terms of increasing their effectiveness. Graphical user interface changes are not enough for a significant increase of detection of review screen anomalies. Current review screens are still not as effective as they should be for the task of voting, since only about half of voters noticed changes.

Everett’s (2007) and Campbell and Byrne’s (2009) results showed that the majority of voters do not notice changes to their ballot selections on a DRE review screen. While this is a significant result, there is a chance it could be due to the fictional ballot used not reflecting real-word issues and the importance of a ballot in an actual election. Acemyan, Kortum, and Payne (2013) investigated this issue by replicating the prior review screen anomaly studies using a ballot with actual candidates and issues that were relevant to the participants. Approximately 75-80% of participants did not notice changes made to their ballot, which is consistent with the results of Everett (2007) and Campbell and Byrne (2009). The real-world relevance of the ballot to the participant did not increase the likelihood of detecting changes in the ballot. The effect of unnoticed change in review screen anomalies has proven to be very robust.

**Alternative Approaches**

Two main tasks must be accomplished to help develop possible solutions for the ineffective DRE review screen design. The first task is to define the current method of review in DREs as well as other possible review methods. The strengths and weaknesses of each review method should then be examined.
Defining Current and Alternative Review Methods

End-of-Ballot. The most commonly used DRE systems generally have end-of-ballot review screens (Everett, 2007). End-of-ballot confirmation (see Figure 3) allows voters to review and confirm all of their selections at once by presenting a screen showing the selections for every race or contest (Everett, 2007).

In-Line. One alternative to end-of-ballot review is an in-line confirmation approach. In-line verification allows voters to confirm their selections immediately after each race, with no global ballot review upon ballot completion. Figure 5 displays a general depiction of in-line confirmation in terms of its use in a DRE.

Figure 5. Diagram depicting the concept of in-line confirmation.

Dual. A combination of in-line and end-of-ballot verification could also be an alternative DRE review method. Voters are able to verify their selections immediately after each race, but are also given a global overview of their selections at the end of the ballot (see Figure 6).
Figure 6. An example of dual confirmation in a DRE.

Strengths and Weaknesses

To gain a thorough understanding of each verification method, their strengths and weaknesses must be addressed. A review of the literature has shown five main topics that could impact the review methods’ usability and performance: efficiency, cognitive demand, “fleeing voter” phenomenon, previous implementation, and redundancy. Each verification method has its own unique strength or weakness in each topic.

**Efficiency.** Efficiency is a prime component of usability in a system (ISO 9241-11, 1998), which has been proven to be important in voting (Everett, 2007; Greene, Byrne, & Everett, 2006). In this case, it is how quickly a voter can complete a ballot on a DRE with each verification method. The general characteristics of each method of review tend to have an effect on their efficiency. A comparison of each of the three verification methods reveals that end-of-ballot review holds the advantage over in-line and dual verification in terms of efficiency.

**End-of-Ballot.** The compact design of end-of-ballot review screens gives a global view of ballot selections. For instance, if a straight-party voter wanted to make sure that they voted for all democratic candidates, an end-of-ballot review screen would allow for that evaluation in an
efficient manner. A voter would simply need scan the review screen to ensure that the political parties are the same throughout. This kind of overall review creates a quick and effective way to verify selections on ballots of varying sizes.

In the event that a voter did not vote straight-party, the end-of-ballot review still affords quick verification of contest selections. Voters can check their selections at once and do not incur an additional time cost that would be associated with a sequential review, where voters confirm their selections after each contest (Holmes, 2013). A voter could simply need to glance over their selections to ensure correctness, instead of taking time to confirm each selection individually.

**In-Line.** The sequential nature of verification for in-line confirmation could lead to longer ballot completion times, negatively affecting efficiency. An increased ballot completion time for an instance of in-line review has been shown in in Holmes (2013). In-line verification effectively doubles the length of the ballot since voters’ must confirm their selection after each contest. For example, a ballot with 27 contests would have the addition of 27 selection confirmations. A voter essentially must make 54 selections, instead of 27 on a ballot.

Although the efficiency could be decreased with an in-line confirmation as shown in Holmes (2013), it does not necessarily mean that performance in other components of usability will decrease as seen in Everett, Greene, Byrne, Wallach, Derr, Sandler, and Torous (2008). Two versions of VoteBox, sequential and direct access, were tested for usability. The sequential version of VoteBox presented the ballot serially, with one contest per page. The direct access version allowed voters to navigate directly to any race. Overall, voters took longer to vote on the sequential DRE than the direct access DRE due to serial presentation of races, similar to in-line confirmation. In terms of effectiveness and satisfaction, the sequential DRE rated higher than the direct access. This result occurred even though the sequential DRE was less efficient than the
direct access DRE. This shows that the overall usability of the system does not necessarily suffer
due to inefficiency.

**Dual.** The dual verification approach exaggerates the projected decreased efficiency
found with in-line confirmation (Holmes, 2013; Everett, et al., 2008). Since the dual review has
both an in-line and end-of-ballot review, the time cost associated with both methods is additive.
This could cause a DRE employing dual verification to become very inefficient and lower its
overall usability.

**Cognitive Demand.** The role of cognitive demand in DRE review could have a
significant effect on performance and usability. A verification method placing less cognitive
demand on a voter should perform better than one that places a heavier cognitive demand on
voters. Several theories of memory and recall should be considered when measuring the amount
of cognitive demand each review method places on voters. The first theory, generally well
known in psychology, is that short term memory is limited (Miller, 1956). It was believed that
this limitation was 7 plus or minus 2 items (Miller, 1956), but has been recently modified to a
capacity of about 4 “chunks” or pieces of information (Cowan, 2001). Another theory of interest
is the list length effect. The list-length effect explains that longer list of items are recalled less
than items on a shorter list (Ratcliff, Clark, & Shiffrin, 1990). The amount of time between
contest selections and review should be also considered. A ubiquitous finding in psychology is
that memory for an item decreases as the time between presentation and recall increases (Howard
& Kahana, 2002). This theory was also supported by Postman and Philips (1965) who found that
shorter times between the presentation of items and their recall result in increased memory for
those items.
In addition to memory and recall, the level of difficulty in detecting erroneous selections among correct ones could also influence the amount of cognitive demand a review method places on voters. Reviewing a ballot for correctness is an example of a signal-detection problem (Green & Swets, 1966). In ballot review, a voter must discriminate between the signal (an erroneous selection or selections) and noise (correct selections). The level of difference between the signal and noise in a review screen represents the level of sensitivity. The greater the sensitivity between the signal and noise, the easier it is for a signal to be correctly discriminated from noise. The main purpose of a review screen is for voters to review their selections for accuracy. Thus, the design of the review screen should make it easy for voters to do so with the highest probability of noticing errors. For example, making undervotes salient by highlighting them on the review screen could help differentiate a possible erroneous selection from a correct one more than if it were not highlighted.

Each review method has its own level of cognitive demand. Based on the described criteria for cognitive demand, in-line verification has the projected advantage over the other review methods.

**End-of-Ballot.** End-of-ballot review may place a high level of cognitive demand on voters. Selker (2007) states that problems recalling previously selected voting choices are significant in an end-of-ballot style review since the typical American ballot election often has more than 500 unique selection possibilities. This task may seem trivial for those who are cognitively adept and are enthusiastic voters, but for those who do not fit those specific categories, the end-of-ballot review screen may be a large undertaking (Selker, 2007).

In 2000, an election in Oregon contained 31 propositions on a ballot (Selb, 2008). In this election, voters would be required to recall how they voted on every single proposition in order
to verify their ballot. In addition to the 31 propositions, voters would also have to correctly recall how they voted for each race. In this instance, ballot length would represent list length in regards to the list-length effect. Therefore, an end-of-ballot review for this scenario could easily have a list length of over 30 items. This may not mean much for the straight-party voter checking their selections for the correct party, but when it comes to remembering selections for a long list of propositions, it could be a problem. Propositions are generally referred to by number on the review screen, and a voter would be at least challenged with the task of remembering what that proposition was and whether or not their vote was represented as intended.

In addition to the possibility of numerous propositions, there is still the chance that a voter is faced with a ballot with many elective offices. For example, the average 2010 Harris County, Texas ballot contained 57 elective offices and the 2010 Cuyahoga County, Ohio ballot contained 35 elective offices (Wheaton, 2013). Both situations would pose a problem for short term memory capacity and list length, especially for those who do not vote straight-party.

Cognitive demand required by end-of-ballot review could also be negatively affected by the amount of time between contest selection and review. For end-of-ballot review, recall occurs after the entire ballot is complete which could be up to several minutes. This time lapse could decrease the chance of accurate recall for selections, leading to errors or malicious changes not always being detected by voters (Howard & Kahana, 2002).

Additionally, several voters in Acemyan, Kortum, and Payne’s (2013) study indicated that they would prefer to review their selections in an in-line fashion due to the difficulty of remembering how they voted in every contest by the time the end-of-ballot review screen appeared (Acemyan, Kortum, & Payne, 2013). This issue may be related to the level of cognitive demand placed on voters when using end-of-ballot review.
Looking at end-of-ballot review as a signal detection problem shows how a seemingly simple cognitive task such as searching for an incorrect selection can become increasingly difficult. The level of sensitivity between signal and noise could vary between different ballots and voters. A straight-party voter may only be interested in what political party they voted for, allowing for a very high level of sensitivity on the review page. Conversely, a voter may not be straight-party, and thus may need to verify candidate names for each contest. In that situation, sensitivity would be lowered, making it more difficult to discern an erroneous selection from the correct ones. It is also important to note that, regardless of the level of sensitivity, the high amount of noise on an end-of-ballot review could negatively affect the probability of a voter correctly finding an error on the review screen (Wickens & Hollands, 1999).

**In-Line.** The cognitive load generated by in-line review is thought to be relatively low due to its direct countering of nearly every cognitive demand aspect shown in end-of-ballot review. In-line confirmation aids voters in error detection because the length of selections to be remembered for confirmation is reduced to one. This would enable voters to easily remember their selection for accurate confirmation as it is well within the bounds of short-term memory and is also a short list length.

In-line verification excels in reducing cognitive demand by nearly eliminating the time between contest selection and review. It allows for the immediate, uninterrupted recall of a voter’s selection. This means that memory accuracy for the selection should be optimal according to prominent findings of recall in psychology (Howard & Kahana, 2002; Postman & Phillips, 1965). In regards to a version of in-line verification, Selker (2007) explains that having selection feedback coincide with selection demands less cognitive resources for error detection.
It helps reduce the memory and matching requirements for voters with cognitive disabilities (Selker, 2007).

In terms of signal detection, in-line verification outperforms other verification methods. This type of verification represents a very simple signal detection problem because of the extremely low noise level. This creates an ideal situation for voters to detect errors or changes to their ballot.

**Dual.** Verifying selection in-line as well as at the end of the ballot encompasses the cognitive demand of both methods, making it the most cognitively demanding method. However, the increased cognitive demand may be countered by the advantage of providing two points of verification. The first point is the in-line confirmation completed while voting, while the second point is the review of all selections at the end of the ballot. Having a multi-step verification system may improve voters’ accuracy, which is significant for system usability (ISO, 9241-11) and election integrity, by possibly increasing the chance for voters to detect errors on their ballot.

The process of multi-step verification has been widely research in the area of medical error prevention. Ghandi et al. (2005) studied the medication dosage error rate associated with chemotherapy ordering process at a cancer institute. The error rate observed in the study was comparable to or lower than error rates found in similar studies according to Ghandi et al. (2005). The authors posited that the lower error rates could possibly be attributed to multiple verification points in the chemotherapy ordering process which were conducted by an attending physician, nurse, and pharmacist. Multi-step verification or “double checking” was also stressed as an important factor in lowing medical error by Cohen, Anderson, Attilio, Green, and Pruemer (1996). The conclusions from both Ghandi et al. (2005) and Cohen et al. (1996) support the
notion that the multiple levels of verification provided by a dual review approach could help lower error rates, increasing the accuracy of the system.

**The “Fleeing Voter” Phenomenon.** The “fleeing voter” phenomenon is a well-documented occurrence in both experimental and actual election settings. It refers to an error occurring when voters complete their ballot but do not click the button to submit it at the review screen (Everett, 2007). This phenomenon is an example of a postcompletion error. Postcompletion errors are an omission of an additional step required for a task after the main goal has been completed is omitted (Byrne & Bovair, 1997). When the main goal of completing a ballot has been accomplished, voters omit the last step of pressing the button to submit their ballot at or after the review screen. This phenomenon is very pertinent to the performance of review screens, since their design can either facilitate or hinder this kind of error.

The “fleeing voter” phenomenon has been documented by Everett (2007), Siegel and Doherty (2006), and Mulligan and Hall (2004). Everett (2007) found that 6% of participants committed this error during the experiment. According to Mulligan and Hall (2004), voters in Texas, Louisiana, Alabama, New Jersey, Pennsylvania, and South Carolina have entered voting booths only to discover that the previous voters had not cast their ballots.

This phenomenon is significant in terms of election integrity because there are no protocols to handle such errors in an election. Some states have rules pertaining to this situation, but most poll workers may not follow these rules consistently due to a lack of training (Everett, 2007). It has been reported that in previous elections some poll workers cast the abandoned ballot, while others cleared the ballot. Also, subsequent voters who are faced with a non-submitted ballot have the opportunity to change that ballot and submit it. They could then fill out
and submit their own ballot, essentially voting twice in the election (Everett, 2007). The consequences of this kind of misconduct and confusion are substantial in elections.

The design of end-of-ballot review screens might facilitate this error, in-line verification could be more resistant to the error.

**End-of-Ballot.** The presentation of the review screen after voters have completed their ballot facilitates the “fleeing voter” phenomenon. The review screen adds an additional step after the main goal of completing the ballot, encouraging the occurrence of a postcompletion error. A review method that enables this type of error could severely affect election integrity due to the aforementioned consequences.

**In-Line.** One way to implement in-line confirmation is for the system to submit the ballot after the last contest selection has been confirmed. This could help prevent the “fleeing voter” phenomenon by removing the any extra steps required to submit the ballot after the main goal of completing the ballot has been accomplished. This removes the criteria necessary for committing a postcompletion error.

**Dual.** The dual verification approach employs end-of-ballot review, therefore making it susceptible to the “fleeing voter” phenomenon. This could be a weakness in this style of verification.

**Previous Implementation.** The concept of previous implementation for review methods is important for two reasons. First, successful implementation of a method provides evidence that it is a viable candidate for DRE review. Secondly, previous implementation could affect users’ familiarity with the method. Having a system or layout that is familiar to users has been shown to
increase subjective usability (Sauro, 2011). In this case, end-of-ballot review ranks above the other review methods.

**End-of-Ballot.** End-of-ballot confirmation has the advantage of familiarity with voters because of its widespread usage in most commercially available DREs (Everett, 2007). Its continued use in DREs for elections could also support its viability as a verification method.

**In-Line.** In-line confirmation has not been explicitly tested in a DRE, but it has been tested in an Interactive Voice Response (IVR) voting system. Holmes (2013) conducted a study assessing the usability of an IVR voting system. This system allowed voters to cast their ballot via a touch-tone telephone. Because this system had a purely auditory interface, the ballot was presented serially. Instead of placing a traditional end-of-ballot review, which would have been difficult due to the auditory interface, the system was designed with an in-line review. This allowed voters to confirm their selections immediately after making them. This system was tested for its usability and compared to other voting methods, specifically the bubble ballot. It was found that the error rate of the IVR voting system was similar to that of the bubble ballot and other traditional voting methods. The design of the IVR voting system, including the in-line review, was posited to be transferable to other types of technology such as the DRE (Holmes, 2013). However, this does not guarantee successful implementation of in-line confirmation on a DRE platform because of the differences in modality between the two interfaces.

In addition to testing the IVR voting system, Holmes (2013) collected survey data on the design of the system. Specifically, the survey asked how much users liked the in-line voting verification and whether or not they preferred it to end-of-ballot verification. A large majority of users (79.3%) indicated that they would prefer to confirm their selections immediately after making the selection and not at the end of the ballot.
A close variation of in-line confirmation was tested in a DRE system in a study by Selker, Rosenweig, and Pandolfo (2006). This system, DRE with voter verified audio audit trail (VVAAT), provided audio feedback to voters other their selections as they made them. Though voters did not have to confirm their selection to move on to the next race, as in the in-line confirmation described for this study, voters were made aware of each selection via auditory response. Voters in Selker, Rosenweig, and Pandolfo’s (2006) study were told how to vote on a majority of the races, with any difference recorded by the system from the way participants were told to vote counting as errors. The accuracy of ballots for the DRE with VVAAT was compared to that of three different voting methods. One method was a DRE with a voter verified paper audit trail (VVPAT). The DRE with VVPAT printed a paper showing voters’ selections once the ballot was completed. Another voting method used was a DRE with contemporaneous VVPAT. This DRE printed voters’ selections as they made them. The last method tested was an optical scan ballot where voters mark their selections on a paper ballot which is scanned by a machine once completed. The DRE with VVAAT produced half the number of selection errors than the DRE with contemporaneous VVPAT. The DRE with VVAAT also had a higher number of voters that noticed errors on their ballots (60% noticed errors) than optical scan ballots (0% noticed errors) and DRE with VVPAT (17% noticed errors). The trend shown in this study could support the value of immediately making voters aware of their selections for each contest.

Dual. In my review of the literature, I have not been able to find any documentation showing the usage of dual verification in a voting system. Though both review methods used in dual verification have been used in some capacity in a voting setting, both methods have not been explicitly tested. Only a version of dual verification was found to be tested in medical technology studies (Ghandi et. all, 2005; Cohen et al., 1996).
**Redundancy.** Users often become desensitized to redundant information, which on a review screen could cause users to ignore and click through verification (Faulkner, 2006; Queensbery, 2001). This could actually facilitate users committing errors instead of preventing them, leading to similar results as in the Everett (20078) and Campbell and Byrne (2009) studies. Thus, it should be avoided in order to optimize the effectiveness of a review screen. In this regard, end-of-ballot review has the advantage over the other alternative review methods.

**End-of-Ballot.** The property of redundancy is not present in end-of-ballot review. Voters are only required to review and confirm their selections once. This could help reduce the issue of desensitization to review.

**In-Line.** The nature of in-line review is inherently redundant. The number of selection reviews and confirmations would be equal to the length of the ballot. The redundant presentation may incite desensitization to the review screen, causing voters to simply click-through the verification without reviewing their selection.

Though redundancy may be a drawback for in-line review, it does offer the property of forcing voters to use view a review screen before progressing through the ballot. A significant predictor of whether or not voters noticed review screen anomalies was the length of time spent on the review screen (Everett, 2007). The longer voters spent on the review screen, the higher the chance of detecting errors. It should be noted that the causal direction of this correlation is unclear. The error(s) could be the cause of increased time spent on the review screen or increased time spent on the review screen could helped improve error detection. With the latter causality in mind, in-line review could encourage voters to spend more time on the review screen, increasing their chance of detecting errors.
**Dual.** Dual verification amplifies the issue of redundancy because not only must voters verify their choice sequentially, but at the end of the ballot as well. Voters may become desensitized to both methods of verification, possibly nullifying the positive attributes of multi-step verification. Subjective usability could also be negatively affected by the exaggerated redundancy of the review method. Though subjective usability may possibly suffer due to the redundancy, studies by Bailey (1993) and Barnum, Henderson, Hood, and Jordan (2004) showed that a system with high performance is not always preferred by users. Purchase, Allder, and Carrington (2001) also stated that user preference does not always correspond with high performance. This may be the case with dual confirmation in that it potentially provides a very accurate system with two-step verification at the cost of perceived usability.

**Effect of Navigation on Review**

In Holmes (2013), the IVR voting system did not permit voters to go back to previous races after their selections had been confirmed. This design was implemented due to certain design constraints of the system. The decision to remove that navigational control from users raised the question of whether navigation style impacts the usability and performance of ballot verification. Many commercially available DREs, as well as those used in studies, have end-of-ballot review and allowed voters to navigate back to races that they have completed while still voting on the ballot (Everett, et al., 2008). This design choice seems to be prevalent, yet the impact of navigation on the review screen has not been explicitly studied.

Cambell, Tossell, Byrne, and Kortum (2001) researched the usability of a custom electronic voting system for mobile phones. This system included the implementation of a button allowing users to navigate back to previous races (back navigation), yet there was no explicit reasoning stating why this design choice was made nor its impact on review screen effectiveness.
Everett et al. (2008) presented a similar situation as the previous study. VoteBox was used to examine the usability and performance of different ballot navigation styles in DREs. While ballot navigation was investigated in the study, the effect of basic ballot navigation on review was not explored.

Furthermore, the impact of navigation on usability has been studied in other circumstances, without the specific understanding of its relevance to ballot review in a DRE. Sandler, Derr, and Wallach (2008) modeled VoteBox’s interface after typical DREs, which in turn are derived from conventional computerized kiosks such ATMs and airline check-in kiosks. This model includes the feature of back navigation, which allows users to go back to previously visited pages, much like the back button on web browsers. This feature is suggested to be implemented on information kiosks to improve usability by Maguire (1999). It is also suggested for use by the US Department of Health and Human Services (United States Department of Health and Human Services, 2006) well as the Voluntary Voting System Guidelines (VVSG) (U.S. Election Assistance Commission, 2012). Both sources voice the suggestion of back navigation to improve usability, but offer no research to support whether it has a positive effect on ballot review in DREs.

Since the literature does not explicitly state why back navigation is recommended, it can only be assumed that it may have stemmed from basic human factors interface design guidelines, such as those found in Nielsen and Molich (1990). Though allowing users to have an “exit” in a system is a well-founded human factors design guideline, there is value in investigating this standard as it pertains to review screens. Exploring this area could help determine which method of navigation enhances the usability and performance of review screens.
Study

The prevalent style of review for DREs has proven to be unsuccessful in helping voters to verify their ballots for accuracy. The repercussions of an ineffective review screen have warranted an investigation of alternative forms of verification in DREs. The absence of knowledge of navigation and its effect on the usability and performance of review screens also pushed for an exploration of this topic.

The objective of this research was to establish baseline performance and usability data for alternative methods of DRE ballot verification as well as determine which method performs best as an error prevention tool. It was also hoped to learn what effects navigation may have on review screen performance and usability. The performance and usability of in-line and dual confirmation methods were compared to end-of-ballot confirmation (as well as each other) to determine whether they varied from the standard form of DRE review.

All methods of review with both forms of navigation were tested via a mock national election using VoteBox Lite. All variables of interest were tested on three usability metrics recommended by the National Institute of Standards and Technology (NIST) (Laskowski, Autry, Cugini, Killam, & Yen, 2004). These metrics are efficiency (time to complete a ballot), effectiveness (accuracy), and satisfaction (subjective usability) as defined by the International Organization for Standardization (ISO 9241-11). The results helped determine whether in-line and dual confirmation were viable review options, in as well as if the ability to navigate backwards through a ballot had any effect on performance.

It is important to note that this study was not a replication of Everett’s (2007) study, as it did not investigate voters’ ability to detect malicious ballot changes. While it may be beneficial to recreate Everett’s (2007) study in order to compare review screen performance, it is more
important to first establish basic performance and usability information for each confirmation
type, assess which method is best, and determine if navigation affects review screen
performance. Once the actual viability of the methods have been determined, further testing of
their usefulness as review screens can begin.

Method

Participants

One hundred one participants were recruited from Rice University’s undergraduate
subject pool. Eight participants were removed from analysis due to having an error rate above
15\% on each review method. This signified non-compliance or a lack of understanding from the
participants. The method of exclusion is similar to those used in previous voting work (Byrne,
Greene, & Everett, 2007). All participants reported having normal or corrected to normal vision.
Participants received course credit for their participation.

The age of participants ranged from 18-23 years, with a mean age of 19.3 years (SD =
1.1). Fifty-two participants were male and forty-one participants were female. Forty-nine (52\%)
participants had previous voting experience with various voting systems. Ninety-one participants
reported their ethnic backgrounds and can be seen in Table 1.
Table 1

*Distribution of Participant Ethnicities*

<table>
<thead>
<tr>
<th>Ethnicity</th>
<th>Frequency</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>African American</td>
<td>4</td>
<td>4.4%</td>
</tr>
<tr>
<td>Asian American/Asian</td>
<td>28</td>
<td>30.8%</td>
</tr>
<tr>
<td>Caucasian</td>
<td>47</td>
<td>51.6%</td>
</tr>
<tr>
<td>Mexican American/Chicano</td>
<td>3</td>
<td>3.3%</td>
</tr>
<tr>
<td>Other Hispanic/Chicano</td>
<td>3</td>
<td>3.3%</td>
</tr>
<tr>
<td>Multiracial</td>
<td>6</td>
<td>6.6%</td>
</tr>
</tbody>
</table>

**Design**

A mixed design was used in this study. There was one between-subjects variable, navigation type. Navigation type had two levels: back navigation and no back navigation, which as the titles suggest dictated whether or not participants could navigate backwards through a ballot. Half of participants were randomly placed in the back navigation condition, while the remaining participants were placed in the no back navigation condition. The within-subjects factor was confirmation type. This variable had three levels: in-line, end-of-ballot, and dual verification. All participants used the custom version of VoteBox to submit their ballot using each of the three verification types.

The dependent variables were the three metrics of usability as defined by ISO 9241-11 (ISO, 1998): efficiency, effectiveness, and satisfaction. Efficiency was measured by the amount of time subjects took to complete the ballot, or total ballot completion time. Effectiveness measured accuracy and was assessed by the number of errors committed while using the system. Errors were determined by comparing the ballot selections made by the user to the voting slate, which directed participants how to vote in each contest. Differences in the selections recorded by
VoteBox from the slate were counted as errors. Error rate was calculated on a per race basis by dividing the number of errors committed by the total number of possible errors for the ballot. The System Usability Scale (SUS) was used to measure satisfaction. In both navigation conditions, the order in which each review system was presented to participants was counterbalanced, presented in every permutation, to prevent any ordering effects that may occur. Each order presented is shown in the list below:

1. In-Line, End-of-Ballot, Dual
2. In-Line, Dual, End-of-Ballot
3. Dual, End-of-Ballot, In-Line
4. Dual, In-Line, End-of-Ballot
5. End-of-Ballot, In-Line, Dual
6. End-of-Ballot, Dual, In-Line

The order in which each subject voted was recorded with the goal of analyzing any effects vote order may have had on the dependent variables, specifically satisfaction. Participants were split equally into each vote order group. System metrics such as time spent on the review screen, selection changes, back navigation usage, and general user activity were recorded and analyzed as well.

Materials

Apparatus. A custom DRE fashioned after VoteBox was used for this experiment. This DRE, designated as VoteBox Lite, lacks the security features included in the original VoteBox and seeks to only replicate its graphical user interface. End-of-ballot, in-line, and dual review methods were implemented on the system.
End-of-ballot review allowed voters to make selections for the entire ballot and review those selections at once at the end. The review screen implemented the same design changes as recommended by Campbell and Byrne (2009). The changes included increasing the salience of undervotes, displaying party information, and emphasizing the importance of ballot review. Users were instructed on the review screen how to alter their selections for a race. Figure 7 is a depiction of the end-of-ballot review screen for VoteBox Lite. The changes implemented in the review screen are seen in Figure 7 as follows: selections highlighted in orange are undervoted contests, candidate parties are located to right of each selection, and the words “*very important*” are located below the ballot review instructional step.

![Review Choices](image)

**Figure 7.** Screenshot of the end-of-ballot review screen in VoteBox Lite.

In-line confirmation required users to confirm their contest selections immediately after making them via a “yes” or “no” prompt (Figure 8). The confirmation buttons were placed in the same location as the navigation buttons, “next” and “previous”, respectively. This means users did not have to move their mouse from the “next” button to confirm their selection on the
confirmation page. If a user did not make a selection, their undervote was highlighted, similar to the end-of-ballot review, and again required verification (Figure 9). In-line confirmation was previously mentioned to have the unique attribute of preventing the fleeing voter phenomenon, achieved by allowing voters to submit their ballot when confirming their last vote, instead of requiring an additional step after verification to submit. Users were alerted with an on screen message that confirming their last vote would submit the ballot. A depiction of this event is shown in Figure 10.

![Figure 8. Screenshot of VoteBox Lite in-line confirmation screen with a contest selection.](image-url)
Figure 9. Screenshot of VoteBox Lite in-line confirmation screen without contest selection.

Figure 10. VoteBox Lite in-line confirmation ballot submission screen.
As the name suggests, dual confirmation included both the in-line and end-of-ballot review. Every detail executed in each review method for VoteBox Lite was included in the dual review method, with the exception of in-line review ballot submission. Users submitted their ballot in the dual verification method in the same manner as end-of-ballot review.

Navigation was presented one of two ways on VoteBox Lite, with or without back navigation. One instance allowed users to navigate to previous races on the ballot (back navigation), while the other instance removed this navigational feature (no back navigation). This feature was included to test whether or not the review screens were affected by navigation type. Eliminating the ability to back navigate only involved removing the “Previous” button from individual contests so that voters could not move to previous races while voting on a ballot. It did not affect any other functionality of VoteBox Lite.

Figure 11. Screenshot of a contest with back navigation button highlighted.
Figure 12. Screenshot of a contest without back navigation button.

VoteBox Lite was created as a shell DRE with the sole purpose of testing alternative review and navigation methods. No security measures were implemented in VoteBox Lite, nor was this system designed for commercial release. This DRE was built for research purposes only. Specific information about the behavior of the system can be found in Appendix A.

Apple iMac desktop computers with 20-inch screens were used to run VoteBox Lite. The ballot used in studies by Everett et al., (2008), Everett (2007), and Campbell and Byrne (2009) was serially presented on the DRE. The ballot contained 21 races at the national, state, county, and non-partisan level and 6 propositions which were taken from various county or state ballots. VoteBox Lite logged ballot completion times and all user interactions with the system. VoteBox Lite required users to enter a pin to begin voting (Figure 13). Standard voting instructions, as
found in VoteBox, were provided on the DRE before participants began voting (see Figure 14).

![Figure 13. Screenshot of VoteBox Lite PIN screen.](image)

![Instructions](image)

**Figure 14.** Screenshot of VoteBox Lite voting instructions screen.

**Other Materials.** Subjective usability of the system was measured using the SUS, which contains ten items answered on a Likert scale (Brooke, 1996). The SUS has the benefit of not
being system specific, so it can be used to assess the subjective usability of many types of systems, including VoteBox Lite. The assessment produces scores that range from 0-100, with 0 being “Not Acceptable” and 100 being “Excellent” (Bangor, Kortum, & Miller, 2009). Several surveys evaluated subjects’ prior voting experiences, demographic information, and opinions about the three review screen types (see Appendix B).

**Procedure**

The experiment began with participants completing a Rice University Institutional Review Board (IRB) approved consent form. Participants then filled out a demographic survey that also collected previous voting history information. Once the survey was completed, subjects were randomly given a voting slate with either a majority of democratic or republican selections and given instructions to vote according to this slate and to save all questions and comments until they were finished with the experiment. The instructions ensured that no experiment related interruptions occurred that could skew the efficiency variable. After instruction, participants completed a ballot for each verification type using VoteBox Lite. Immediately after each ballot was submitted, participants completed a SUS assessment regarding only the review method they just used (explicit instruction was given). Lastly, the survey pertaining to the review screen design of VoteBox was given after participants completed a ballot for all three review types.

**Results**

A two-way mixed ANOVA was used for all statistical analyses, unless otherwise noted. Huynh-Feldt adjustment was used on $p$-values when $\varepsilon < .85$. Pairwise $t$-tests with Bonferroni correction were used to further understand any differences found between groups.
Usability Results

Efficiency. Figure 15 displays the mean ballot completion time as a function of verification type. A reliable effect of verification type was found, $F(2, 182) = 5.59, p = .004$, $MSE = 1,371, \eta^2_p = .06$. There was no evidence supporting a significant effect of navigation type on ballot completion time, $F(1, 91) = 1.90, p = .17, MSE = 2,099, \eta^2_p = .02$. An interaction between navigation and verification type was not supported by the data, $F(2, 182) = 2.37, p = .10, MSE = 1,371, \eta^2_p = .03$.

Figure 16 shows a comparison of efficiency of VoteBox Lite with other voting methods, including VoteBox, from previous studies by Everett et al. (2008), Campbell et al. (2010), and Greene et al. (2006). It must be noted that different populations assessed in this comparison were different, therefore a direct comparison cannot be made.

![Figure 15. Mean ballot completion time in seconds by verification type.](image-url)
Pairwise $t$-tests revealed that the efficiency for end-of-ballot verification was significantly different from that of dual confirmation, but not from in-line confirmation (end-of-ballot and dual: $t(92) = 3.13, p < .01$; in-line and end-of-ballot: $t(92) = 1.42, p = .48$). No significant difference in efficiency was found for in-line and dual confirmation (in-line and dual: $t(92) = 2.06, p = .13$).

**Effectiveness.** No reliable effects of verification were found for error rate, $F(1.4, 123.5) = 2.08, p = .14, MSE = .73, \eta^2_p = .02$ (Huynh-Feldt adjustment used). Error rates for each verification type can be seen in Figure 17. Navigation type did not have a reliable effect on effectiveness, $F(1, 91) = 1.16, p = .28, MSE = 1.81, \eta^2_p = .01$. No evidence supporting an interaction between navigation and verification type was found, $F(1.4, 123.5) = 1.68, p = .20, MSE = .73, \eta^2_p = .02$ (Huynh-Feldt adjustment used).
Figure 17. Mean error rate for each verification type. Error bars represent standard error.

Figure 18. Mean error rate comparison of VoteBox Lite with other voting methods.
Two types of errors could have been made by participants during the experiment:

- **Omission errors**: no selection made when the intent was to make a selection.
- **Wrong choice errors**: the selection made is not the one that was intended.

A third type of error, *overvote*, could not have occurred in this study. Overvote errors can occur in two ways: a voter selects more candidates than allowed in one race or a voter makes a selection in a race that they intended to omit (extra vote error). VoteBox Lite did not allow voters to vote for more than the allowed number of candidates. The slate did not have any omitted races, therefore voters could not make an extra vote error. Wrong-choice errors were the only errors made by participants in this experiment. Errors were also calculated on a by ballot basis which distinguishes ballots that have at least one error with those that do not.

Table 2

*Frequency of Errors by Ballot*

<table>
<thead>
<tr>
<th></th>
<th>No Errors</th>
<th>Errors</th>
</tr>
</thead>
<tbody>
<tr>
<td>In-Line</td>
<td>92</td>
<td>1</td>
</tr>
<tr>
<td>End-of-Ballot</td>
<td>90</td>
<td>3</td>
</tr>
<tr>
<td>Dual</td>
<td>88</td>
<td>5</td>
</tr>
</tbody>
</table>

**Satisfaction.** The mean SUS scores for each verification method are located in Figure 19. Data supported a reliable effect of verification type on satisfaction, $F(2, 182) = 11.71$, $p < .001$, $MSE = 56.45$, $\eta^2_p = .11$. There was no significant effect of navigation type, $F(1, 91) = 1.76$, $p = .19$, $MSE = 203.28$, $\eta^2_p = .02$. There was minimal evidence supporting an interaction between navigation and verification type, $F(1.36, 123.45) = 1.68$, $p = .20$, $MSE = .73$, $\eta^2_p = .02$. A comparison of mean SUS scores for the VoteBox Lite system against other voting methods can be seen in Figure 20.
Figure 19. Mean SUS scores for each verification method.

Figure 20. Mean SUS scores for VoteBox Lite and other voting methods.
Pairwise $t$-tests were used with Bonferroni correction that modified the p-value to examine any specific differences between the verification methods. The tests revealed that end-of-ballot verification scores were reliably different from in-line and dual verification (end-of-ballot and in-line: $t(92) = 5.47, p < .01$; end-of-ballot and dual: $t(92) = 2.51, p = .04$). No significant differences were found between in-line and dual confirmation SUS scores (in-line and dual: $t(92) = 2.03, p = .14$).

Review Screen Time

The amount of time spent on the review screens for each verification method was analyzed. End-of-ballot review time consisted of the time participants spent on the review screen displayed at the end of the ballot. For in-line verification, review time was defined as the sum of time spent on each confirmation screen. Review time for dual verification was calculated by combining the total in-line confirmation screen time with the end-of-ballot review screen time. Figure 21 shows the mean review screen time for each verification method. A significant effect of verification type was found on the time spent on review screens, $F(2, 182) = 27.17, p < .001, MSE = 522.99, \eta^2_p = .23$. An effect of navigation was not found, $F(1, 91) = 1.05, p = .31, MSE = 738.36, \eta^2_p = .01$. There was also no convincing evidence supporting an interaction between navigation and verification, $F(2, 182) = .41, p = .67, MSE = 522.99, \eta^2_p < .01$. Pairwise $t$-tests with altered p-values Bonferroni correction indicated a reliable difference between all levels of verification (in-line and end-of-ballot: $t(92) = 4.49, p < .01$; in-line and dual: $t(92) = 3.91, p < .01$; end-of-ballot and dual: $t(92) = 6.15, p < .01$).
Figure 21. Mean time spent on the review screen for each verification method.

The review screen time for dual verification was broken into two components, in-line review and end-of-ballot review, so that a direct comparison with the other respective verification methods could be made. There was insufficient evidence to support any differences between the split dual review times and their respective counterparts. The mean review time for each race with in-line verification and the in-line portion of dual verification is shown in Figure 23 and Figure 24 respectively.
Figure 22. Review times for in-line and end-of-ballot verification methods compared to their respective counterparts for dual verification.

Figure 23. Mean review time for each race with in-line verification.
Figure 24. Mean review time for each race with the in-line confirmation portion of dual verification.

Vote Order

The order in which participants used each verification method did not have a significant effect on any variable:

- **Efficiency:** $F(1, 87) = 1.69, p = .15, MSE = 2.042, \eta^2_p = .09$
- **Effectiveness:** $F(1, 87) = .73, p = .61, MSE = 1.85, \eta^2_p = .04$
- **Satisfaction:** $F(1, 87) = .70, p = .63, MSE = 208.47, \eta^2_p = .04$
- **Review Screen Time:** $F(1, 87) = 1.21, p = .31, MSE = 730.39, \eta^2_p = .07$

There was a reliable interaction between vote order and verification for efficiency and satisfaction variables (Efficiency: $F(10, 174) = 14.62, p < .01, MSE = 799.56, \eta^2_p = .46$;
Satisfaction: $F(10, 174) = 4.08, p < .01, MSE = 48.14, \eta^2_p = .20$. The interactions are represented below in Figure 25 and Figure 26.

Table 3

*Efficiency Interaction Means*

<table>
<thead>
<tr>
<th></th>
<th>Order 1</th>
<th>Order 2</th>
<th>Order 3</th>
<th>Order 4</th>
<th>Order 5</th>
<th>Order 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>In-Line</td>
<td>164.98</td>
<td>169.25</td>
<td>118.31</td>
<td>129.26</td>
<td>139.31</td>
<td>115.78</td>
</tr>
<tr>
<td>End-of-Ballot</td>
<td>108.77</td>
<td>121.56</td>
<td>122.99</td>
<td>116.81</td>
<td>166.06</td>
<td>153.39</td>
</tr>
<tr>
<td>Dual</td>
<td>122.11</td>
<td>172.78</td>
<td>164.14</td>
<td>183.85</td>
<td>129.28</td>
<td>133.50</td>
</tr>
</tbody>
</table>

*Figure 25.* Interaction between verification type and vote order on efficiency.
Table 4

Satisfaction Interaction Means

<table>
<thead>
<tr>
<th></th>
<th>Order 1</th>
<th>Order 2</th>
<th>Order 3</th>
<th>Order 4</th>
<th>Order 5</th>
<th>Order 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>In-Line</td>
<td>88.28</td>
<td>87.81</td>
<td>82.86</td>
<td>91.00</td>
<td>83.75</td>
<td>83.59</td>
</tr>
<tr>
<td>End-of-Ballot</td>
<td>90.63</td>
<td>89.38</td>
<td>86.96</td>
<td>95.00</td>
<td>92.66</td>
<td>94.22</td>
</tr>
<tr>
<td>Dual</td>
<td>85.16</td>
<td>91.72</td>
<td>92.68</td>
<td>89.33</td>
<td>89.69</td>
<td>82.66</td>
</tr>
</tbody>
</table>

*Figure 26. Interaction between verification and vote order on satisfaction.*

The interactions were investigated using simple main effects. There was a reliable effect of verification on efficiency at every level of vote order. This lead to further analysis of the data by plotting ballot completion time by the order of trial, collapsing across verification type (Figure 27). This allowed for a visual representation of the average amount of time taken to complete the first, second, and third ballot, regardless of verification type. Figure 27 revealed a learning effect,
which may have decreased the time to complete a ballot with each trial. The learning effect in combination with the effect of verification on efficiency could be the driving force behind the interaction.

Satisfaction data revealed a reliable effect of verification on orders 3, 4, 5, and 6 (refer to Methods for order definitions), but not order 1 and 2. Simply stated, when in-line verification was the first method used, then verification type may have had no bearing on how participants rated a review method. Alternatively, verification type did have an effect on the SUS score of a review method when dual or end-of-ballot review was presented first.

- **Order 3:** $F(2, 30) = 4.90, p = .02, \text{MSE} = 69.47, \eta^2_p = .27$
- **Order 4:** $F(2, 30) = 4.19, p = .03, \text{MSE} = 30.35, \eta^2_p = .23$
- **Order 5:** $F(2, 30) = 8.43, p < .01, \text{MSE} = 39.04, \eta^2_p = .36$
- **Order 6:** $F(2, 30) = 7.71, p < .01, \text{MSE} = 85.59, \eta^2_p = .34$

![Figure 27. Time to complete a ballot for first, second, and third trial.](image-url)
Verification Method Survey

The majority of participants (52%) chose end-of-ballot as their preferred method in the experiment, with the majority (56%) choosing in-line confirmation as the least preferred method. Dual confirmation was stated to be the best review method for aiding error prevention (60%) as well as being best suited for voting (63%). The majority of participants strongly agreed that the time to vote using in-line (40%) and end-of-ballot (69%) verification was acceptable. Dual confirmation rendered a majority response (37%) of “slightly agree” in regards to completing the ballot in an acceptable amount of time. A large majority of participants (71%) believed that accuracy was more important than efficiency on a ballot. Specific survey results can be found in Appendix C.

Discussion

Current DRE review screen design has been proven to be unreliable in aiding voters in detecting errors on their ballots. The current method of basic navigation, though widely used, lacks research in the value of its use in DREs. In response to these issues, this study attempted to gather baseline performance and usability information for alternative methods of review and navigation as well as to determine the viability of these alternatives. The results of this study could have large implications in not only the design of DREs but other processes where verification is critical.

Usability

The alternative methods were tested on the three metrics of usability (efficiency, effectiveness, and satisfaction), as prescribed by ISO 9241 part11 (ISO 9241-11, 1998). These metrics were examined and compared to determine the standard of usability performance for the alternatives as well as ascertain their merit as compared to the current DRE design conventions.
**Efficiency.** End-of-ballot confirmation was more efficient than dual confirmation. The longer ballot completion time for dual confirmation compared to end-of-ballot could possibly be attributed to the additional time cost of confirming selections in-line. As stated before, in-line confirmation doubles the length of the ballot in a way, since each ballot contest must be confirmed on a separate page in the DRE. In contrast, end-of-ballot verification’s attribute of global review could have facilitated the faster ballot completion times.

An investigation of the interaction between verification type and voting order on efficiency revealed that there could be two main driving forces behind the effect: the main effect of efficiency and the learning effect (see Figure 27). Participants completed their ballots slowest on their first voting trial and fastest on their last voting trial, regardless of which verification method was used. For example, end-of-ballot verification was the fastest method for every vote order except those where it was the first method used. The main effect of verification type would generally show that end-of-ballot verification is the fastest method, but this is not the case when the learning effect is involved.

While efficiency is an important factor in usability, voting is an area where effectiveness may take precedence over it. Participants stated in the survey that when it comes to voting, efficiency is a worthy tradeoff for accuracy. Even though participants were willing to make this time sacrifice, from a global standpoint, using either review alternative may not sacrifice efficiency. Participants were able to complete a ballot with dual confirmation in an average time of two minutes and thirty seconds and those using in-line confirmation in an average time of two-minutes and twenty seconds. These times are consistent with those of other voting methods in use today, which means that either alternative could be viewed as an acceptable for voting in
terms of efficiency. However, this does not support the use of the alternative review methods over the currently used end-of-ballot review.

**Effectiveness.** No evidence of differences between the error rates for each verification type was found. This suggests that the alternative review methods perform as well as the end-of-ballot review, but not better. This result provides no evidence supporting the notion that in-line or dual verification improve error detection or prevention.

In-line verification had only one incidence of error on a ballot. This may speak to its efficacy as a review screen. Dual verification had the highest number of by ballot errors, with 5 ballots containing at least one error. The high incidence of errors on the dual verification ballots occurred contrary to the believed advantage that multi-step verification would offer in preventing errors. Fatigue, stated to be a possible weakness of dual verification, may seem to explain the amount of by ballot errors committed, but Figure 23 and Figure 24 show that this was not the case. The figures show a steep decline in the time spent on the review screen for in-line verification and the in-line confirmation portion of dual verification after the first race, instead of a steady decline as would be expected if fatigue were occurring. Dual verification was selected by users as the method requiring the most effort in terms of memory and attention, so that could have been the cause, in part, of the higher number of by ballot errors.

Evidence supporting an effect of vote order on error rate was not found. This could mean that the order in which participants voted with each verification type had no bearing on the number of by errors made. All verification methods, regardless of by ballot errors performed similarly to other voting systems.

**Satisfaction.** End-of-ballot review had a reliably higher average SUS score than dual and in-line review. This could be due to the redundancy of the dual confirmation, as predicted. Even
with the lower SUS scores, dual and in-line confirmation were still rated as “Excellent” systems and are deemed “acceptable” (Bangor, Kortum, & Miller, 2009). Also, all review methods’ scores seem to be at the same level as those of other voting systems. With that said, this result does not support the change of end-of-ballot review as the prominent review method in DREs as it is also rated as an “Excellent” and “acceptable” system and had the highest SUS score of the alternative review methods.

The interaction between verification type and vote order on satisfaction affected how participants scored each verification method only when in-line verification was not the first method used. Participants generally rated the first method used the highest with the exception of the fourth voting order. The ratings for the second and third methods used seemed to vary depending on the vote order. In some cases, the methods were rated in an order that corresponded with their order of use. For example, if end-of-ballot verification was used second, it was rated second highest.

**Performance and Design**

Navigation type did not reliably influence any performance or usability metric. It may be likely that whether or not voters are able back navigate does not interact with the performance or usability of review screens. However, 29% of participants utilized the back navigation feature. This alone could stand as reasoning to keep the feature, as it does not reduce or change system performance. Allowing back navigation also keeps DREs analogous with other voting methods such as the paper ballot, where voters are able to look back at previous races while voting.

The evidence did not suggest that the order in which participants voted affected how they performed with each verification type. However, vote order in combination with verification type
did affect efficiency and satisfaction variables. Any further testing within-subjects testing of verification methods may need to account for ordering effects.

In Everett (2007), time spent on the review screen was a large predictor of whether or not voters noticed changes on their ballot. The longer voters spent on the review screen, the greater the chance they would notice ballot changes. Though, as previously discussed (page 22), the direction of causality for this correlation is unknown. In one aspect of the relationship between error detection and time spent on the review screen it is possible that a verification method that encourages longer time spent on the review screen could aid voters in detecting errors. All verification methods reliably differed from each other with respect to time spent on review screens. This suggests that dual confirmation elicited the longest time spent on the review screen (47.9 seconds) while end of ballot brought about the least time spent on the review screen (17.1 seconds). Those using in-line confirmation spent an average time of 29.6 seconds on the review screen. In-line confirmation nearly doubles the review screen time of end-of-ballot confirmation, while dual confirmation nearly triples the review screen time. While both alternatives resulted in longer time spent on the review screen, it is important to understand that this does not necessarily mean that participants were attentive to the review screens. A participant may be on the review screen, but not actually actively reviewing their selections. Therefore, the purpose of the review screen is rendered ineffective no matter how long participants spend on the screen.

The fleeing voter phenomenon seen in Everett (2007) and multiple other studies and reports did not occur in this study. The in-line confirmation prevented this error by submitting the ballot once the last contest was confirmed. Dual and end-of-ballot confirmation both required participants to submit their ballot after the review screen was presented, but the submit ballot button was made very salient by centering its location and making its background green instead
of grey like the other navigational buttons. The submit button was also provided on the review screen as opposed to after it. The absence of the fleeing voter phenomenon may also have been due to the sample being Rice University undergraduates, inducing a best case scenario for the experiment. As a result, this study provided no evidence supporting any verification method above another in regards to preventing the fleeing voter phenomenon.

**Qualitative Results**

The survey results gave insight to participants’ subjective views of each verification method beyond that of statistical data. Survey results indicated that overall, participants felt that dual confirmation was the most appropriate verification method for voting, even though they rated it as their least preferred experience in the experiment. Participant feedback seems to show that users are aware of the importance of verification, despite the quality of experience associated with it. End-of-ballot review was rated by users to be the method requiring the least demand on memory and attention. This view is contradictory to the hypothesis of it being the most cognitively demanding method. This result could be due to the active nature of review that dual and in-line confirmation require, or because of users’ familiarity with the review type.

Dual confirmation was also stated to be the best method to notice malicious changes to the ballot. It was also thought by participants to be the best method when it comes to activities related to online banking. This result could indicate that users value multiple points of verification in situations where accuracy is paramount. Overall, the efficiency of all methods was declared acceptable. However, participants indicated that for voting, accuracy is more important than efficiency. This could mean that users would be more willing to take longer to vote, as long as their vote was accurate. In general, the qualitative results in this study show that dual verification is the preferred method when accuracy is crucial.
The results of this study did not support the replacement of end-of-ballot review with in-line or dual verification. The data imply that dual and in-line review are no better than the current standard of end-of-ballot review (see Table 5). All review methods performed consistently with other voting systems (traditional and non-traditional), indicating that the alternative verification methods may be viable options for DRE. However, the data still support the continued use of end-of-ballot review. Overall, the evidence does not encourage changing the current standard of DRE review, but baseline performance data has been established for future comparisons. As for ballot navigation in DREs, the evidence was inconclusive as to whether or not navigation type provided any advantages or disadvantages. The data did reveal that the back navigation feature was used by nearly 30% of participants, suggesting it has importance in a DRE.

Table 5

*Verification Strengths and Weaknesses*

<table>
<thead>
<tr>
<th></th>
<th>End-of-Ballot</th>
<th>In-Line</th>
<th>Dual</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Efficiency</strong></td>
<td>Advantage</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Effectiveness</strong></td>
<td>No Difference</td>
<td>No Difference</td>
<td>No Difference</td>
</tr>
<tr>
<td><strong>Satisfaction</strong></td>
<td>Advantage</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Review Time</strong></td>
<td></td>
<td></td>
<td>Advantage</td>
</tr>
<tr>
<td><strong>Subjective Voting Preference</strong></td>
<td></td>
<td></td>
<td>Advantage</td>
</tr>
</tbody>
</table>

**Limitations**

One of the largest limitations of this study was the population selected for this experiment. The participants consisted solely of Rice University undergraduates who are all
relatively young (19.3 years), with normal or corrected to normal vision, and no cognitive impairments. Using this sample could only represent the best case scenario for testing the alternative methods, which may not accurately reflect how the methods may perform in an actual voting situation. This limitation should also be considered when examining the information that compared the usability metrics of VoteBox Lite with those of traditional voting methods, which was gathered using the general voting population. Diversifying the sample by using the general voting population would eliminate any potential bias.

This study did not investigate how the alternatives performed with straight party ballots or with contests requiring more than one selection. Both situations can occur in actual voting conditions, and they should be tested with the alternative methods to ensure a complete understanding of review and navigation performance under real voting situations.

Participants did not vote in an environment similar to that in a national election. Privacy dividers were not used, nor were participants required to stand while using the system. These differences could possibly have affected the efficiency of the verification methods as well as some of the qualitative results. Although there is no direct evidence of this, if participants were required to stand while voting, an accelerated effect of fatigue could occur. Voters could become uncomfortable standing while voting, leading to them rushing through the ballot and ignoring the review screen(s). The lack of privacy dividers could lead to participants feeling that their responses were not kept confidential from other participants who could be looking on to their screen. Therefore, a participant might rush through the ballot to reduce the chance of other participants viewing their selections.
Future Directions

This dissertation provides performance information for alternative review methods. The results showed the efficacy of the alternatives, but did not reveal which alternatives were best in improving notice of malicious changes and errors in a ballot. The current study could act as a foundation for further investigation into which method works best as an error prevention tool. Now that the alternatives have been proven acceptable, additional testing with them can occur.

A replication of Everett’s (2007) study would be very useful in determining which verification method performs best as a review screen. The introduction of known errors could provide more evidence in establishing, if any, which review method is actually superior.

Investigating these review methods in areas other than voting, such as online banking or stock trading could prove to be useful. The results in this study could be vastly different when participants are faced with different tasks. For the task of confirming a transfer of funds, in-line verification may prove to be the best method, since the task is much shorter than completing a ballot. The differences in efficiency between the verification methods could very well shrink. Different situations could cause differences in the usability of review methods. The fact that this study did not support the change of the standard method of review does not mean that the same would be true when testing alternative review methods in other tasks.

Conclusion

Studies have demonstrated that current DRE review screens are not optimal in aiding voters in detecting errors. This study explored the usability and viability of alternative methods of review in DREs in order to determine the best method for error prevention and provided baseline performance data for these review methods. The study also investigated the effect of navigation on review screen performance and usability. The alternative methods of ballot
verification (in-line and dual) were tested for usability and baseline performance data in a custom DRE system called VoteBox Lite. The tested methods proved to be viable alternatives in terms of usability and performance. However, no evidence was found that the alternative methods of review were better or worse than end-of-ballot review, suggesting that the current standard of review should remain in place. Qualitative results supported the use of dual confirmation in voting over in-line and end-of-ballot confirmation. The baseline performance and usability data in this experiment can now be used for comparison to other DRE design alternatives as well as other voting methods. Further research should include testing the alternatives with malicious ballot changes to provide more evidence for determining which alternative is best suited as an error prevention tool. These alternative should also be tested outside the area of voting, as the implications of this study could be generalized to other electronic platforms where verification is critical.
References


Holmes, D. (2013). *Vote-By-Phone: Usability of an IVR voting system with adjustable audio speed.* (Master's Thesis). Rice University, Houston, TX.


http://www.measuringusability.com/blog/prior-exposure.php


http://ww2.gazette.net/stories/090606/montcou181637_31942.shtml


Appendix A

VoteBox Lite Design

No Back Navigation

In-line Confirmation

1. Ballot Navigation

2. Ballot Submission

Diagram showing the flow of events for Ballot Navigation and Ballot Submission.
No Back Navigation

End-of-Ballot Confirmation

1. Ballot Navigation

1. Race is presented
   - Race 1: A, B, C

2. User makes selection
   - Race 1: A

3. User continues through ballot
   - Race 2: A

4. User makes selection
   - Race 2: C

5. User completes last contest
   - Race 27: C

6. Review screen

2. Review and Ballot Submission

1. Review screen
   - Race 1: B
   - Race 2: B
   - Race 27: A

2. User selects race to change
   - Race 1: C

3. User is taken to selected race
   - Race 1: B

4. User changes selection
   - Race 1: C

5. User is returned to updated review screen
   - Race 1: C
   - Race 2: B
   - Race 27: A

Submit
No Back Navigation

Dual Confirmation

1. Ballot Navigation

2. Ballot Review

3. Ballot Submission
Back Navigation

In-line Confirmation

1. Ballot Navigation

- If user back navigates to a previous race, they will not have to re-confirm when navigating back to the latest race if there are no changes to the selections.

2. Ballot Submission
Back Navigation

End-of-Ballot Confirmation

1. Navigation

<table>
<thead>
<tr>
<th>Race 1</th>
<th>Race 1</th>
<th>Race 2</th>
<th>Race 2</th>
<th>Race 27</th>
<th>Review</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>B/A</td>
<td>Race 1: B</td>
</tr>
<tr>
<td>B</td>
<td>B</td>
<td>B</td>
<td>B/C</td>
<td>B</td>
<td>Race 2: B</td>
</tr>
<tr>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>Race 27: A</td>
</tr>
</tbody>
</table>

1. Race is presented
2. User makes selection
3. User presses "Previous"
4. User makes selection
5. User makes selection on last race
6. Review screen

2. Ballot review and confirmation

<table>
<thead>
<tr>
<th>Review</th>
<th>Review</th>
<th>Race 1</th>
<th>Race 1</th>
<th>Review</th>
</tr>
</thead>
<tbody>
<tr>
<td>Race 1: B</td>
<td>Race 1: B</td>
<td>A</td>
<td>A</td>
<td>Race 1: C</td>
</tr>
<tr>
<td>Race 2: B</td>
<td>Race 2: B</td>
<td>B</td>
<td>B</td>
<td>Race 2: B</td>
</tr>
<tr>
<td>Race 27: A</td>
<td>Race 27: A</td>
<td>C</td>
<td>C</td>
<td>Race 27: A</td>
</tr>
</tbody>
</table>

1. Review screen
2. User selects race to change
3. User is taken to selected race
4. User changes selection
5. User is returned to updated review screen
Back Navigation

Dual Confirmation

1. Navigation

![Diagram of navigation process]

- If user back navigates to a previous race, they will not have to re-confirm when navigating back to the latest race if there are no changes to the selections.

2. Ballot Review

![Diagram of ballot review process]
VoteBox Lite Screenshots

**Review Choices**

Down below are the selections you have made. Click on a race to change your selection. To submit your ballot, click the Submit button, and your vote will be recorded.

- President and Vice President: Leonard Conley, Jo Melody, and Howard Fiske
- Senate: Tom Amundson
- Representatives: Congress: Robert Harkins
- Governor: Don Howard
- Lieutenant Governor: Sharon Van Pelt
- Attorney General: Bob Ryan

**Presidential and Vice Presidential Candidates**

- Donald Trump
- Hillary Clinton
- Gary Johnson
- Jill Stein

**Propositions of State Importance**

- Proposition 1: Yes
- Proposition 2: Yes
- Proposition 3: Yes
- Proposition 4: Yes
- Proposition 5: Yes
- Proposition 6: Yes

President and Vice President

You voted for:

Gordon Bearce and Nathan Maclean
Republican

Is this correct?

[No] [Yes]
Representative in Congress

You did not make a selection

Is this correct?

---

Proposition 6

Once you confirm your selection your ballot WILL BE SUBMITTED. You will not be able to go back and change any other selections

You voted:

Yes

Is this correct?
President and Vice President

To make your choice, click on the box next to the candidate's name. A checkmark will appear next to your choice. If you want to change your choice, just click on a different candidate's box. Please select only ONE candidate.

- Gordon Bawng and Nathan Maclean
- Vernon Stanwyk Albury and Richard Rigby
- Janette Froman and Chris Aponte

---

President and Vice President

To make your choice, click on the box next to the candidate's name. A checkmark will appear next to your choice. If you want to change your choice, just click on a different candidate's box. Please select only ONE candidate.

- Gordon Bawng and Nathan Maclean
- Vernon Stanwyk Albury and Richard Rigby
- Janette Froman and Chris Aponte
Welcome. You are about to begin voting in a mock election at Rice University. There are three steps to voting in this election.

**Step One:** Finish reading these instructions.
**Step Two:** Make your choices. This is where you will choose your candidates and vote for or against the propositions. You will make one choice per page.
**Step Three:** Review your choices. This is where you will be able to see all of the choices you have made.
**Step Four:** Record your vote by clicking on the ‘Submit’ button. This is the last step. Once you complete this step, your vote will be recorded.

When you get to each of these three steps, you will see more detailed instructions. At the bottom of the screen, there will be buttons you can click with your mouse. Click the ‘Begin’ button to start voting.
Appendix B

Background and General Voting Survey

1. Age: _______

2. Gender: ____ Male    ____ Female

3. Year or Position at Rice University (if applicable):

   _____ Freshman
   _____ Sophomore
   _____ Junior
   _____ Senior
   _____ 5th +
   _____ Graduate
   _____ Staff
   _____ Faculty
   _____ Other/None

4. If you are a student, in what division(s) is/are your planned major(s)? Check all that apply.

   _____ Humanities
   _____ Social Sciences
   _____ Natural Sciences
   _____ Engineering
   _____ Architecture
   _____ Music

5. Do you have normal or corrected to normal vision? _____ No    _____ Yes

6. Do you consider yourself to have a reading disability? _____ No _____ Yes

7. Are you left or right handed? _____ Right    _____ Left    _____ Ambidextrous

8. Are you a native English speaker? _____ No_____ Yes

   If no, what is your native language? _________________
9. Can you touch type? (Can you type without looking at the keys?)
   ____ No   ____ Yes

10. How many hours per week do you use a computer?
    ____ less than 5 hours
    ____ between 5 and 10 hours
    ____ between 10 and 20 hours
    ____ between 20 and 30 hours
    ____ between 30 and 40 hours
    ____ over 40 hours

11. Please rate your level of computer expertise (1 = novice, 10 = expert)
    1 2 3 4 5 6 7 8 9 10

12. Which of these activities do you use a computer for? Check all that apply.
    ____ Word Processing (e.g. Microsoft Word)
    ____ Programming (e.g. Java, C++, Scheme)
    ____ Web design
    ____ Graphic Design (e.g. Adobe Photoshop, Illustrator)
    ____ Video Editing
    ____ Personal Finance (e.g. Quicken, Turbo Tax)
    ____ Games
    ____ Music
    ____ Multimedia (e.g. encyclopedias; interactive CDs)
    ____ Spreadsheet management (e.g. Microsoft Excel)
    ____ Data Analysis (e.g. SAS, SPSS)

13. What is your political affiliation?
    ____ Republican
    ____ Democrat
    ____ Libertarian
    ____ Independent
    ____ Other, please specify: ______________________

14. How many national elections have you voted in?
    ____ 0 to 5
    ____ 6 to 10
    ____ 11 to 15
    ____ 15 to 20
    ____ 21 or more
15. In which state(s) and county(s) have you voted in a national election?
__________________________________________________________________

16. How many other elections of any type (local, school board, etc.) have you voted in?
   ____ 0 to 5
   ____ 6 to 10
   ____ 11 to 15
   ____ 15 to 20
   ____ 21 or more

17. In which state(s) and county(s) have you voted in other types of elections?
__________________________________________________________________

***For questions 18 - 25, please answer keeping in mind your previous voting experience in any type of election (not including voting you did in this study). If you have never voted, please skip questions 18 - 25.***

18. Do you typically cast your vote on an absentee ballot?
   ____ No ____ Yes

19. Please indicate how many times you have used each type of technology or ballot to cast your vote in any election.
   ____ Fill in the bubble (or box)
   ____ Connect the arrows (or lines)
   ____ Open response
   ____ Lever machines
   ____ Punch cards
   ____ Electronic – touch screen
   ____ Electronic – other
   ____ Don’t know
   ____ Other, please specify: __________
20. Have you ever felt worried about figuring out how to use the ballot or technology to cast your vote?

_____ No  _____ Yes

21. Have you ever felt that time pressure caused you to rush, make a mistake, or leave a choice blank when you would not otherwise have done so?

_____ No  _____ Yes

22. If you have felt time pressure, did this prevent you from voting?

_____ No  _____ Yes

23. Do you typically vote a straight-party ticket?

_____ No  _____ Yes

24. Do you typically cast a vote for every office on the ballot?

_____ No  _____ Yes

25. When you voted in an election, have you ever been unsure if your vote was cast correctly or would be counted?

_____ No  _____ Yes

If yes, please describe the situation:

26. Have you been following the news about computer voting and potential security concerns?
(Please choose one)

_____ No, not at all
_____ Yes, somewhat
_____ Yes, very closely

27. Has news about computer voting and potential security concerns affected your trust of these systems?

_____ No  _____ Yes
Why or why not?

__________________________________________________
__________________________________________________
__________________________________________________

28. How often do you use an ATM (Automated Teller Machine) to get money or complete other transactions at a bank, grocery store, or other location?
   ____ never
   ____ very infrequently
   ____ occasionally (for example 1-4 times a year)
   ____ often (for example once a month)
   ____ frequently (for example once a week or more)

29. What is your current occupation? ____________________________________________

30. Please indicate the highest level of education you have completed.
   ____ Some high school
   ____ High school or G.E.D.
   ____ Some college or Associate's degree
   ____ Bachelor’s degree or equivalent
   ____ Postgraduate degree (such as M.A., Ph.D., M.D., J.D.)

31. Are you:
   ____ African American
   ____ American Indian
   ____ Asian American
   ____ Caucasian
   ____ Mexican American or Chicano
   ____ Other Hispanic or Latino (please specify) ________________________________
   ____ Multiracial (please specify) ________________________________
   ____ Other (please specify) ________________________________
Confirmation Method Survey

Definitions

In-line confirmation (confirm candidate selections immediately after making the selection)

End-of-Ballot confirmation (confirm all races at the end of the ballot)

Dual Confirmation (confirm each selection immediately after selection AND confirm all races at the end of the ballot)

1. Rate your preference of each confirmation method on a scale of 1-3 (1 is most preferred, 3 is least preferred)
   a. In-line confirmation _______
   b. End-of-Ballot confirmation _______
   c. Dual Confirmation _______

2. Which method of confirmation required the least amount of effort (memory, attention, etc.)?
   a. In-line confirmation
   b. End-of-Ballot confirmation
   c. Dual Confirmation

3. Which method of confirmation do you feel would allow you make the fewest mistakes?
   a. In-line confirmation
   b. End-of-Ballot confirmation
   c. Dual Confirmation

4. Which method of confirmation do you feel would allow you to make the most mistakes?
   a. In-line confirmation
   b. End-of-Ballot confirmation
   c. Dual Confirmation
5. Imagine that your voting machine had been tampered with and changed your candidate selections for 3 random races (without your knowledge). The only way for you to notice these changes would be using the ballot confirmations. In your opinion, which confirmation method would be the best in helping you detect the changes in your ballot?
   a. In-line confirmation
   b. End-of-Ballot confirmation
   c. Dual Confirmation

6. In regards to voting, which confirmation method do you feel is most appropriate?
   a. In-line confirmation
   b. End-of-Ballot confirmation
   c. Dual Confirmation

7. In regards to online banking (i.e. money transfers), which confirmation method do you feel is most appropriate?
   a. In-line confirmation (confirm each step in the money transfer process immediately after each step)
   b. End-of-Ballot confirmation (confirm all steps in the money transfer process at the end)
   c. Dual Confirmation (confirm all money transfer steps immediately after each step AND at the end of the process)

8. The amount of time taken to complete a ballot using the in-line confirmation method was acceptable.
   a. Strongly disagree
   b. Slightly disagree
   c. Neither disagree or agree
   d. Slightly agree
   e. Strongly agree

9. The amount of time taken to complete a ballot using the end-of-ballot confirmation method was acceptable.
   a. Strongly disagree
   b. Slightly disagree
   c. Neither disagree or agree
   d. Slightly agree
   e. Strongly agree

10. The amount of time taken to complete a ballot using the dual confirmation method was acceptable.
a. Strongly disagree
b. Slightly disagree
c. Neither disagree or agree
d. Slightly agree
e. Strongly agree

11. I believe accuracy is more important than time in regards to voting.
a. Strongly disagree
b. Slightly disagree
c. Neither disagree or agree
d. Slightly agree
e. Strongly agree
Ballot

GENERAL ELECTION BALLOT
HARRIS COUNTY, TEXAS
NOVEMBER 4, 2009

* TO VOTE, COMPLETELY FILL IN THE OVAL ☐ NEXT TO YOUR CHOICE.
  * Use only the marking device provided or a number 2 pencil.
  * If you make a mistake, do not hesitate to ask for a new ballot. If you erase or make other marks, your vote may not count.

<table>
<thead>
<tr>
<th>STRAIGHT PARTY VOTING</th>
<th>STATE</th>
<th>COUNTY</th>
</tr>
</thead>
<tbody>
<tr>
<td>☐ Republican REP</td>
<td>ATTORNEY GENERAL (Vote for One)</td>
<td>☐ Corey Bohnke REP</td>
</tr>
<tr>
<td>☐ Democratic DEM</td>
<td>RICK ORGAN (Vote for One)</td>
<td>☐ Jennifer A. Laundow DEM</td>
</tr>
<tr>
<td>☐ Gordon Bence REP</td>
<td>COMPTROLLER OF PUBLIC ACCOUNTS (Vote for One)</td>
<td>☐ Dean Coffey REP</td>
</tr>
<tr>
<td>☐ Nathan Maclean DEM</td>
<td>☐ Therese Gutin IND</td>
<td>☐ Gordon Kallas DEM</td>
</tr>
<tr>
<td>☐ Vernon Stanley Albry</td>
<td>☐ Greg Converse DEM</td>
<td>☐ Stanley Saar GP</td>
</tr>
<tr>
<td>☐ Richard Rigby LIB</td>
<td>COMMISSIONER OF GENERAL LAND OFFICE (Vote for One)</td>
<td>☐ Sheryl Ind. LIB</td>
</tr>
<tr>
<td>☐ Janette Froman LIB</td>
<td>RAILROAD COMMISSIONER (Vote for One)</td>
<td>☐ Howard Gardy IND</td>
</tr>
<tr>
<td>☐ Chris Aponte DEM</td>
<td>☐ Polly Rylander REP</td>
<td>☐ Randy H. Ciemens CON</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CONGRESSIONAL</th>
<th>UNITED STATES SENATOR (Vote for One)</th>
<th>COUNTY TAX ASSESSOR (Vote for One)</th>
</tr>
</thead>
<tbody>
<tr>
<td>☐ Cecile Cadieux REP</td>
<td>COMMISSIONER OF AGRICULTURE (Vote for One)</td>
<td>☐ Howard Grady IND</td>
</tr>
<tr>
<td>☐ Fern Brezinski DEM</td>
<td>☐ Polly Rylander REP</td>
<td>☐ Howard Grady IND</td>
</tr>
<tr>
<td>☐ Coree Dery IND</td>
<td>☐ Robert Aron DEM</td>
<td>☐ Howard Grady IND</td>
</tr>
<tr>
<td>☐ Representative in Congress (Vote for One)</td>
<td>☐ Raul H. Rios DEM</td>
<td>☐ Howard Grady IND</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>STATE</th>
<th>GOVERNOR (Vote for One)</th>
<th>COUNTY JUDGE (Vote for One)</th>
</tr>
</thead>
<tbody>
<tr>
<td>☐ Pedro Brouse REP</td>
<td>☐ Jillian Bates REP</td>
<td>☐ Dan Atchley</td>
</tr>
<tr>
<td>☐ Robert Metier DEM</td>
<td>☐ Zachary Millican DEM</td>
<td>☐ Lewis Shulz</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>STATE</th>
<th>STATE SENATOR (Vote for One)</th>
<th>COUNTY JUDGE (Vote for One)</th>
</tr>
</thead>
<tbody>
<tr>
<td>☐ Ricardo Nigh REP</td>
<td>☐ Wesley Steven Millette DEM</td>
<td>☐ Dan Atchley</td>
</tr>
<tr>
<td>☐ Rick Stilley REP</td>
<td>☐ State Representative District 194 (Vote for One)</td>
<td>☐ Lewis Shulz</td>
</tr>
<tr>
<td>☐ Maurice humble IND</td>
<td>☐ State Representative District 2 (Vote for One)</td>
<td>☐ Dan Atchley</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>LIEUTENANT GOVERNOR</th>
<th>JUDGE</th>
<th>COUNTY JUDGE (Vote for One)</th>
</tr>
</thead>
<tbody>
<tr>
<td>☐ Shane Terrill REP</td>
<td>☐ Petra Bencomo REP</td>
<td>☐ Dan Atchley</td>
</tr>
<tr>
<td>☐ Jessie Principe DEM</td>
<td>☐ Susanne Rael DEM</td>
<td>☐ Lewis Shulz</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>STATE</th>
<th>MEMBER STATE BOARD OF EDUCATION DISTRICT 2 (Vote for One)</th>
<th>COUNTY JUDGE (Vote for One)</th>
</tr>
</thead>
<tbody>
<tr>
<td>☐ Peter Vergas REP</td>
<td>☐ Mark Barber DEM</td>
<td>☐ Dan Atchley</td>
</tr>
<tr>
<td>☐ Tim Grady DEM</td>
<td>☐ Presiding Judge Texas Supreme Court Place 9 (Vote for One)</td>
<td>☐ Lewis Shulz</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>PROPOSITIONS</th>
<th>PROPOSITION 1</th>
<th>COUNTY JUDGE (Vote for One)</th>
</tr>
</thead>
<tbody>
<tr>
<td>☐ Yes</td>
<td>☐ No</td>
<td>☐ Dan Atchley</td>
</tr>
</tbody>
</table>

VOTE BOTH SIDES OF BALLOT
<table>
<thead>
<tr>
<th>PROPOSITION 2</th>
<th>PROPOSITION 4</th>
<th>PROPOSITION 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shall the Charter of Harris County be amended to authorize the City Council to review and approve certain intergovernmental agreements and revenue contracts entered into by the City; to permit the City Council to establish its meeting schedule by ordinance; to clarify the circumstances in which the City Council may act by ordinance or resolution; to permit the City Council to adopt by ordinance procedures for the formation and administration of special assessment districts; to permit excused absences of council members for reasons other than sickness; and to make other conforming amendments related thereto in order to eliminate redundant or obsolete provisions of the charter?</td>
<td>Shall there be an amendment to the Texas revised statutes concerning renewable energy standards for large providers of retail electric service, and, in connection therewith, defining eligible renewable energy resources to include solar, wind, geothermal, small hydroelectricity, and hydrogen fuel cells; requiring that a percentage of retail electricity sales be derived from renewable sources, beginning with 3% in the year 2011 and increasing to 10% by 2021; requiring utilities to offer consumers a rebate of $0.01 per watt and other incentives for solar electric generation; providing incentives for utilities to invest in renewable energy resources that provide net economic benefits to customers; limiting the rate impact of renewable energy resources to 50 cents per month for residential customers; requiring public utilities commission rules to establish major aspects of the measure; prohibiting utilities from using condemnation or eminent domain to acquire land for generating facilities not in compliance with the standards; requiring utilities with requirements contracts to address shortfalls from the standards; and specifying election procedures by which the customers of a utility may opt out of the requirements of the amendment?</td>
<td>Shall the Charter of Harris County concerning the powers of the City Council be amended in regard to the sale of city-owned property, to require Council approval for the sale of personal property valued at $500,000 or more, and to clarify language requiring Council approval of any sale of real property?</td>
</tr>
<tr>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>NO</td>
<td>NO</td>
<td>NO</td>
</tr>
</tbody>
</table>

**PROPOSITION 3**

Shall there be an amendment to the Texas constitution concerning recovery of damages relating to construction of real property improvements and, in connection therewith, prohibiting laws that limit or impair a property owner's right to recover damages caused by a failure to construct an improvement in a good and workmanlike manner; defining "good and workmanlike manner" to include construction that is suitable for its intended purposes; and permitting exceptions for laws that limit punitive damages, afford governmental immunity, or impose time limits of specified minimum lengths on filing lawsuits?

| YES |
| NO |

**PROPOSITION 5**

Shall there be an amendment to the Texas constitution concerning election day voter registration, and, in connection therewith, allowing an eligible citizen to register and vote on any day that a vote may be cast in any election beginning on January 1, 2011; specifying election day voter registration locations; specifying that an eligible citizen who registers to vote on election day shall register in person and present a current and valid Texas driver's license or state identification card or other approved documentation; and directing the Texas general assembly, in implementing election day voter registration, to adopt necessary protections against election fraud?

| YES |
| NO |
Republican Slate

President And Vice President:
Gordon Bearce (R)

United States Senator:
Cecile Cadieux (R)

Representative in Congress:
Pedro Brouse (R)

Governor:
Glen Travis Lozier (R)

Lieutenant Governor:
Shane Terrio (R)

Attorney General:
Tim Speight (R)

Comptroller of Public Accounts:
Greg Converse (D)

Commissioner of General Land Office:
Sam Saddler (R)

Commissioner of Agriculture:
Roberto Aron (D)

Railroad Commissioner:
Jillian Balas (R)

State Senator:
Ricardo Nigro (R)

State Representative District 134:
Petra Bencomo (R)

Member State Board of Education
District 2:
Peter Varga (R)

Presiding Judge Texas Supreme Court Place 3:
Tim Grasty (D)
Presiding Judge Court of Criminal Appeals Place 2:
Dan Plouffe (R)
**District Attorney:**
Corey Behnke (R)

**County Treasurer:**
Dean Caffee (R)

**Sheriff:**
Jason Valle (LIB)

**County Tax Assessor:**
Randy H. Clemons (CON)

**Justice of the Peace:**
Deborah Kamps

**County Judge:**
Dan Atchley

**Proposition 1:**
No

**Proposition 2:**
Yes

**Proposition 3:**
No

**Proposition 4:**
Yes

**Proposition 5:**
Yes

**Proposition 6:**
No
Democratic Slate

**President And Vice President:**
Vernon Stanley Albury (D)

**United States Senator:**
Fern Brzezinski (D)

**Representative in Congress:**
Robert Mettler (D)

**Governor:**
Rick Stickles (D)

**Lieutenant Governor:**
Shane Terrio (R)

**Attorney General:**
Rick Organ (D)

**Comptroller of Public Accounts:**
Greg Converse (D)

**Commissioner of General Land Office:**
Sam Saddler (R)

**Commissioner of Agriculture:**
Roberto Aron (D)

**Railroad Commissioner:**
Zachary Minick (D)

**State Senator:**
Wesley Steven Millette (D)

**State Representative District 134:**
Susanne Rael (D)

**Member State Board of Education**
**District 2:**
Mark Baber (D)

**Presiding Judge Texas Supreme Court Place 3:**
Tim Grasty (D)

**Presiding Judge Court of Criminal Appeals Place 2:**
Dan Plouffe (R)
District Attorney:
Jennifer A. Lundeed (D)

County Treasurer:
Gordon Kallas (D)

Sheriff:
Jason Valle (L)

County Tax Assessor:
Randy H. Clemons (CON)

Justice of the Peace:
Clyde Gayton Jr.

County Judge:
Lewis Shine

Proposition 1:
Yes

Proposition 2:
No

Proposition 3:
Yes

Proposition 4:
No

Proposition 5:
Yes

Proposition 6:
Yes
Appendix C

Survey Results

1. Rate your preference of each confirmation method on a scale of 1-3 (1 is most preferred, 3 is least preferred)

<table>
<thead>
<tr>
<th>Method</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>In-Line</td>
<td>12.80%</td>
<td>31.40%</td>
<td>55.80%</td>
</tr>
<tr>
<td>End-of-Ballot</td>
<td>51.60%</td>
<td>27.50%</td>
<td>20.90%</td>
</tr>
<tr>
<td>Dual</td>
<td>40.20%</td>
<td>40.20%</td>
<td>19.50%</td>
</tr>
</tbody>
</table>

2. Which method of confirmation required the least amount of effort (memory, attention, etc.)?

<table>
<thead>
<tr>
<th>Method</th>
<th>In-Line</th>
<th>End-of-Ballot</th>
<th>Dual</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>20.70%</td>
<td>64.10%</td>
<td>15.20%</td>
</tr>
</tbody>
</table>

3. Which method of confirmation do you feel would allow you make the fewest mistakes?

<table>
<thead>
<tr>
<th>Method</th>
<th>In-Line</th>
<th>End-of-Ballot</th>
<th>Dual</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>9.80%</td>
<td>7.60%</td>
<td>82.60%</td>
</tr>
</tbody>
</table>

4. Which method of confirmation do you feel would allow you to make the most mistakes?

<table>
<thead>
<tr>
<th>Method</th>
<th>In-Line</th>
<th>End-of-Ballot</th>
<th>Dual</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>38.00%</td>
<td>56.50%</td>
<td>5.40%</td>
</tr>
</tbody>
</table>

5. Imagine that your voting machine had been tampered with and changed your candidate selections for 3 random races (without your knowledge). The only way for you to notice these changes would be using the ballot confirmations. In your opinion, which confirmation method would be the best in helping you detect the changes in your ballot?

<table>
<thead>
<tr>
<th>Method</th>
<th>In-Line</th>
<th>End-of-Ballot</th>
<th>Dual</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>15.20%</td>
<td>25.00%</td>
<td>59.80%</td>
</tr>
</tbody>
</table>

6. In regards to voting, which confirmation method do you feel is most appropriate?

<table>
<thead>
<tr>
<th>Method</th>
<th>In-Line</th>
<th>End-of-Ballot</th>
<th>Dual</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>10.90%</td>
<td>26.10%</td>
<td>63.00%</td>
</tr>
</tbody>
</table>
7. In regards to online banking (i.e. money transfers), which confirmation method do you feel is most appropriate?

<table>
<thead>
<tr>
<th>In-Line</th>
<th>End-of-Ballot</th>
<th>Dual</th>
</tr>
</thead>
<tbody>
<tr>
<td>15.20%</td>
<td>10.90%</td>
<td>73.90%</td>
</tr>
</tbody>
</table>

8. The amount of time taken to complete a ballot using the in-line confirmation method was acceptable.

<table>
<thead>
<tr>
<th>Strongly Disagree</th>
<th>Slightly Disagree</th>
<th>Neither Disagree or Agree</th>
<th>Slightly Agree</th>
<th>Strongly Agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.30%</td>
<td>8.70%</td>
<td>10.90%</td>
<td>37.00%</td>
<td>40.20%</td>
</tr>
</tbody>
</table>

9. The amount of time taken to complete a ballot using the end-of-ballot confirmation method was acceptable.

<table>
<thead>
<tr>
<th>Strongly Disagree</th>
<th>Slightly Disagree</th>
<th>Neither Disagree or Agree</th>
<th>Slightly Agree</th>
<th>Strongly Agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.30%</td>
<td>3.30%</td>
<td>8.70%</td>
<td>16.30%</td>
<td>68.50%</td>
</tr>
</tbody>
</table>

10. The amount of time taken to complete a ballot using the dual confirmation method was acceptable.

<table>
<thead>
<tr>
<th>Strongly Disagree</th>
<th>Slightly Disagree</th>
<th>Neither Disagree or Agree</th>
<th>Slightly Agree</th>
<th>Strongly Agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.00%</td>
<td>18.50%</td>
<td>13.00%</td>
<td>37.00%</td>
<td>31.50%</td>
</tr>
</tbody>
</table>

11. I believe accuracy is more important than time in regards to voting.

<table>
<thead>
<tr>
<th>Strongly Disagree</th>
<th>Slightly Disagree</th>
<th>Neither Disagree or Agree</th>
<th>Slightly Agree</th>
<th>Strongly Agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.10%</td>
<td>4.30%</td>
<td>5.40%</td>
<td>17.40%</td>
<td>71.10%</td>
</tr>
</tbody>
</table>