RICE UNIVERSITY

System Usability and User Mental Models of Three Verifiable End-to-end Voting Methods: Helios, Prêt à Voter, and Scantegrity II

by

Claudia Ziegler Acemyan

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

Doctor of Philosophy

APPROVED, THESIS COMMITTEE:

Philip Kortum, Chair
Assistant Professor, Psychology

Michael D. Byrne
Professor, Psychology and Computer Science

David Lane
Associate Professor, Psychology, Statistics, and Management

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

HOUSTON, TEXAS
July 2014
ABSTRACT

System Usability and User Mental Models of Three Verifiable, End-to-end Voting Methods: Helios, Prêt à Voter, and Scantegrity II

by

Claudia Ziegler Acemyan

There are many ways voting systems can be maliciously attacked so that election outcomes are altered. In response, voting security experts developed end-to-end (e2e), verifiable voting methods. These systems were intended to be secure, accurate, reliable, and transparent, while still preserving voter anonymity. What is not clear is if these complex, novel systems, which allow voters to check on their ballots after voting, will be usable by every voter. If voting methods are unusable, negative ramifications like disenfranchisement and altered election outcomes could occur. For this reason, system usability and voter mental models of e2e systems must be understood. To address this lacuna in voting research, three e2e methods representative of voter verifiable technologies were studied: Helios, Prêt à Voter, and Scantegrity II.

Three studies were conducted. In the first study, baseline usability data were collected. By having participants vote with each system in a mock election, it was found that the systems were difficult to use. Only 58% of voters were able to cast a ballot, and fewer were able to verify their votes. Also in the first study, the behavioral errors that led
to ballot casting and vote verification failures were identified, along with the potential contributing system design deficiencies. This part of the study revealed that a few design details were driving most of the observed failures. In the second study, voters’ mental models for each voting system were explored. The data supported the claim that voters did not have comprehensive mental models accounting for how the systems work; rather their models emphasized how-to-vote procedures, which were not always correct. In the third study, it was asked if voters even wanted to use the verification systems, and if they did, what form of verification they would expect. Sixty-five percent of voters indicated that they would be interested in checking that their ballot was cast. As for the preferred form of verification, there was not a consensus—indicating that a diverse set of expectations will need to be accounted for when developing the systems. In conclusion, the tested e2e systems were not easily usable by voters, fully understood by them, or in a form that voters might have expected. Yet many of the system problems observed can be fixed, and voters seem to support the idea of auditable voting systems—meaning future effort should be spent improving upon the next generations of e2e systems so that all voters can use secure and transparent voting systems in elections.
DEDICATION

For my husband, Alexandre S. Acemyan, and father, Edward R. Ziegler, who have supported me in so many ways through years of graduate school. Because of your support, I will be able to do work that I love.
ACKNOWLEDGMENTS

My quest to complete a Ph.D. in human factors / HCI psychology has been incredibly rewarding. I have so many people to acknowledge for getting me to this point in my academic career.

I have Philip Kortum to thank for being a wonderful advisor and mentor. Not only did he train me to become a good research scientist, but also he successfully guided me through my doctoral program in three calendar years. In the process I had fun, learned a great deal, and became even more passionate about my research. I really appreciate all of the time he put into talking with me about the field and conducting research. I am extremely fortunate to have had the opportunity to study under his direction.

Mike Byrne and David Lane made me very comfortable with statistics. Thanks to them, I know I will always be able to analyze my data in order to support my research findings. Mike and David also rounded-out my education in human factors, human computer interaction, and engineering psychology as they each have a unique perspective and are great instructors. During this research project, I am especially grateful for all of the time Mike put into helping me develop the studies and giving me honest, pointed feedback.

Dan Wallach helped me understand the basics of e2e, auditable voting systems, assisted me in selecting the three systems tested in this research, and answered numerous questions that arose while I was developing the system prototypes. I really wanted this research to be relevant to and respected by security researchers and auditing experts. Because of his input, I think it will be.
I would like to thank Peter Ryan for his input while I developed the version of PaV tested in this study. Both he and Thea Peacock spent a considerable amount of their time communicating with me via skype and email. I also appreciate that the City of Tokoma Park, Maryland and Alex Essex answered several of my questions about Scantegrity II.

Thanks to Bob Stein I was able to collect all of the data for the vote verification study. I also found the voting class that he co-taught with Mike Byrne and Dan Wallach to be influential on my research.

Other faculty at Rice who impacted my development as a psychologist included Fred Oswald, Randi Martin, and Jim Pomerantz. I want to especially express my gratitude to Fred who spent a substantial amount of time working with me and Phil on developing a trust in voting (TIV) research instrument. While I did not use it in this dissertation research, developing and then using the TIV in other studies helped me to gain insight into how e2e systems differed from one another with respect to trust.

I want to thank my HF/HCI colleagues Chad Tossell, Danae Holmes, Clayton Stanley, Melissa Gallagher, Jeff Zemla, Nicole Howie, Sebastian Thomas, Bryan Campbell, and Jill Piner. Every single one of them offered input on my research at some point in my career at Rice, and they were all incredibly positive and supportive when we presented our research to one another at the weekly HF/HCI seminars. I also must acknowledge that I had several, talented undergraduate researchers help me with this project: David Payne, Jennifer Shafer, Molly Ahn, Mitchel Massey, and Zoe Rogers.

Lanita Martin and Jennifer Gucwa—our department’s graduate student coordinators—made sure that I only had to worry about my work. No matter how big or
small a problem, they always found a solution. Then Lemlem Terke always made certain that I had a nice office in which to work, which definitely made my time at Rice nicer.

From UCLA where I received my M.A., I want to thank Diane Favro, Dana Cuff, and Michael Osman for helping me with my writing. I also want to thank Norton M. Wise, Soraya de Chadarevian, and Jean Louis Cohen for supporting me when I decided to leave architecture and return to psychology.

At Rice University where I received my M.Arch., Carlos Jimenez, John Casbarian, Nana Last, Robert Mangurian, and Mary Ann Ray taught me that I could solve any design problem, no matter how challenging. I now approach problems in HF with the same perseverance.

At the University of Michigan where I completed my B.A. in psychology, Theresa M. Lee and David Kaplan encouraged me to pursue both psychology and architecture. I find that many of the skills that I used as an architect are incredibly helpful when engaged in iterative development and testing of systems. I am glad they emphasized that I could combine both of my interests in some form.

I want to acknowledge my immediate family: Alexandre S. Acemyan, Catherine Eloise “Ouisie” Acemyan, Cynthia Chase, Edward R. Ziegler, Rini Ziegler, Christina Jackson, Macy Jackson, Bruce Jackson, Danielle Wu Baldauff, Nathan Baldauff, Jack Baldauff, and Alvin Lee. My family has always prioritized education and learning, especially my mother and father, which influenced my decision to pursue graduate degrees. Ouisie, my 11-month old daughter, deserves a special thank you for sleeping through the night since she has been two-weeks old (no, I did not let her cry it out). That made life as a graduate student and new parent more pleasant. My husband Alex has also
been an amazing mate and really has been there for me throughout all of life’s adventures that have been thrown at us. Each and every day, Alex has been supportive. He is also a wonderful father who equally shares in the responsibilities of caring for an infant—meaning I was able to simultaneously be a mother and further my career. Even my dogs Poinsettia and Le Corbusier helped out by being welcome distractions at the end of long workdays.

Alex Lara, Ana Rojas, Maria Ceja, Marie Grace, Nelly Martinez, Claudia Herrera, and Karen Gatten must be thanked for providing excellent care for Ouisie while I worked. If I did not have a safe, nurturing place for Ouisie to spend her days at, I would not have finished my Ph.D. by this time.

This research has been supported in part by NIST grant 60NANB12D249 and NSF award CNS-1049723. Any opinions, findings, conclusions, and recommendations expressed in this material are those of the author and do not necessarily reflect the views of the National Institute of Standards and Technology or the National Science Foundation.
TABLE OF CONTENTS

DEDICATION.............................................................................................................. IV

ACKNOWLEDGMENTS .............................................................................................. V

TABLE OF CONTENTS .............................................................................................. IX

LIST OF FIGURES ..................................................................................................... XIV

LIST OF TABLES ....................................................................................................... XVII

LIST OF ABBREVIATIONS ....................................................................................... XVIII

CHAPTER 1 ................................................................................................................. 1

INTRODUCTION

E2E VOTING SYSTEMS .............................................................................................. 1

Helios ..................................................................................................................... 5

Prêt à Voter ............................................................................................................ 11

Scantegrity II ........................................................................................................ 17

RESEARCH PROJECT DESCRIPTION ..................................................................... 23

Study 1—Part 1 of 2 ............................................................................................. 29

Study 1—Part 2 of 2 ............................................................................................. 30

Study 2 ............................................................................................................... 30

Study 3 ............................................................................................................... 31
CHAPTER 2 ........................................................................................................................................ 33

USABILITY OF VOTER VERIFIABLE END-TO-END VOTING SYSTEMS: BASELINE
DATA FOR HELIOS, PRÊT À VOTER, AND SCANTEGRITY II

INTRODUCTION .................................................................................................................................. 33
Understanding the Usability of e2e Voting Systems ....................................................................... 34

METHODS......................................................................................................................................... 36
Participants ......................................................................................................................................... 36
Design ................................................................................................................................................. 37
Procedures .......................................................................................................................................... 39
Materials ............................................................................................................................................ 41

RESULTS .......................................................................................................................................... 43
Vote Casting ......................................................................................................................................... 43
  Effectiveness ....................................................................................................................................... 43
  Efficiency ............................................................................................................................................ 45
  Satisfaction ......................................................................................................................................... 47
Vote Verification .................................................................................................................................... 48
  Effectiveness ....................................................................................................................................... 48
  Efficiency ............................................................................................................................................ 49
  Satisfaction ......................................................................................................................................... 50

DISCUSSION ...................................................................................................................................... 52
  Why are these systems failing? ........................................................................................................... 53
  Where do we go from here? ................................................................................................................ 57

CONCLUSION .................................................................................................................................... 58
CHAPTER 3 ............................................................................................................................................. 59

FROM ERROR TO ERROR: WHY VOTERS COULD NOT CAST A BALLOT AND VERIFY THEIR VOTE

INTRODUCTION ........................................................................................................................................ 59

METHODS ............................................................................................................................................. 61

RESULTS .............................................................................................................................................. 61

Helios ................................................................................................................................................... 62

Vote Casting ......................................................................................................................................... 62

Vote Verification ................................................................................................................................... 66

Prêt à Voter .......................................................................................................................................... 68

Vote Casting ......................................................................................................................................... 68

Vote Verification ................................................................................................................................... 72

Scantegrity II ......................................................................................................................................... 73

Vote Casting ......................................................................................................................................... 73

Vote Verification ................................................................................................................................... 73

DISCUSSION AND CONCLUSION ........................................................................................................ 77

CHAPTER 4 ........................................................................................................................................... 82

USERS’ MENTAL MODELS FOR END-TO-END VOTING SYSTEMS................................................. 82

INTRODUCTION ..................................................................................................................................... 82

METHODS ............................................................................................................................................. 87

Participants .......................................................................................................................................... 87

Design .................................................................................................................................................. 87

Procedures .......................................................................................................................................... 88
RESULTS ......................................................................................................................... 89

System Security and Accuracy ................................................................................... 93
Unnecessary Steps and Procedures ............................................................................. 99
Procedural Uncertainty ................................................................................................ 101
System Preferences ..................................................................................................... 102

DISCUSSION AND CONCLUSION ............................................................................ 104

CHAPTER 5 .................................................................................................................. 109

TO VERIFY OR NOT TO VERIFY, THAT IS THE QUESTION: VOTERS’
ANTICIPATED USE AND EXPECTATIONS OF VOTING VERIFICATION METHODS

INTRODUCTION ............................................................................................................. 109

METHODS .................................................................................................................... 113

Participants ................................................................................................................... 113
Design ............................................................................................................................ 113
Procedure ....................................................................................................................... 115

RESULTS .................................................................................................................... 115

Confidence in System Accuracy .................................................................................... 115
Desire to Verify .............................................................................................................. 116
Verification Format ....................................................................................................... 117

DISCUSSION AND CONCLUSION .............................................................................. 119

CHAPTER 6 .................................................................................................................. 124

GENERAL DISCUSSION & CONCLUSION ................................................................. 124

OVERVIEW OF RESEARCH ......................................................................................... 124

WHY AREN’T THESE SYSTEMS MORE USABLE? ....................................................... 126
MOVING FORWARD WITH THE DEVELOPMENT OF E²E VOTING SYSTEMS............................... 128
FUTURE RESEARCH ............................................................................................................ 131
CONCLUSION .................................................................................................................... 134

REFERENCES ..................................................................................................................... 135

APPENDIX A ...................................................................................................................... 146

HELIOS SYSTEM AND STUDY MATERIALS

APPENDIX B ...................................................................................................................... 151

PAV SYSTEM AND STUDY MATERIALS

APPENDIX C ...................................................................................................................... 155

SCANTTEGRITY II SYSTEM AND STUDY MATERIALS

APPENDIX D ...................................................................................................................... 159

REPUBLICAN DIRECTED VOTING SLATE

DEMOCRATIC DIRECTED VOTING SLATE

APPENDIX E ...................................................................................................................... 161

SYSTEM USABILITY SCALE (SUS) ADAPTED FOR VOTING

SUS ADAPTED FOR VOTE VERIFICATION

APPENDIX F ...................................................................................................................... 163

FINAL PARTICIPANT SURVEY
LIST OF FIGURES

Figure 1. Screenshot of Helios Voting Booth instructions ........................................ 6
Figure 3. Screenshot of the Helios Review Screen.................................................. 8
Figure 4. Screenshot of the page that tells voters they must log in to verify their eligibility.......................................................... 9
Figure 5. Screenshot of the Helios cast vote confirmation page.............................. 9
Figure 6. Screenshot of the voter’s archived ballot .............................................. 10
Figure 7. Screenshot of Helios’ Voters and Ballot Tracking Center ...................... 11
Figure 8. Voting instructions for PaV .................................................................. 12
Figure 9. Card 1/8 of the PaV study ballot ...................................................... 13
Figure 10. Photograph of a voter casting their PaV ballot.................................... 14
Figure 11. An example of a PaV voter receipt...................................................... 15
Figure 12. Screenshot of study’s PaV’s vote verification home webpage............. 16
Figure 13. Screenshot of PaV’s vote validation webpage developed for the research study ............................................................................ 17
Figure 14. Front of the Scantegrity II ballot ..................................................... 18
Figure 15. Scantegrity II vote verification sheet.................................................. 19
Figure 16. Photograph of a completed Scantegrity II ballot, with invisible ink confirmation codes revealed .................................................. 20
Figure 17. Screenshot of Scantegrity II’s vote verification home page developed for this study ............................................................................ 22
Figure 18. Screenshot of Scantegrity II’s cast ballot confirmation page developed for this research project ................................................................. 23

Figure 19. Percentage of cast ballots as a function of voting system, with different colored bars representing perceived and actual cast vote ........................................ 44

Figure 20. Mean per-contest error rate percentage as a function of voting system type, with error bars depicting the standard error of the mean ........................................ 45

Figure 21. Vote casting completion time as a function of voting system, with dashed horizontal lines depicting the mean and solid horizontal lines depicting the median .................................................................................................................................. 46

Figure 23. Percentage of verified votes as a function of voting system, with different colored bars representing perceived and actual verified votes .......................... 49

Figure 24. Verification completion time as a function of voting system, with dashed horizontal lines depicting the mean and solid horizontal lines depicting the median .................................................................................................................................. 50

Figure 25. SUS rating for the vote verification process as a function of voting system, with dashed horizontal lines depicting the mean and solid horizontal lines depicting the median .................................................................................................................................. 51

Figure 26. Helios screen that voters see after encrypting their ballot but before logging in so that they can later cast the ballot ...................................................................................... 64

Figure 27. Photograph of the scanner used in this research project ................................. 71

Figure 28. An incorrectly completed Scantegrity II ballot and vote confirmation sheet. 76

Figure 29. Two instances of participants’ drawn mental models for Helios ....................... 90

Figure 30. Two instances of participants’ drawn mental models for PaV ............................ 91
Figure 31. Two instances of participants’ drawn mental models for Scantegrity II........ 92

Figure 32. Frequency as a function of voting system type for the system participants
indicated was most secure.............................................................................................. 98

Figure 33. Frequency as a function of voting system type for the system participants
indicated was least secure ............................................................................................... 99

Figure 34. Frequency as a function of voting system type for the system participants
would most prefer to use in a real election ................................................................. 103

Figure 35. Frequency as a function of voting system type for the system participants
would least prefer to use in a real election ................................................................. 103

Figure 36. Screenshots of Helios’ (shown left), PAV’s (shown center), and Scantegrity
II’s (shown right) vote verification pages that may be viewed by voters if they
choose to check on their cast ballots. The PAV and Scantegrity II pages are
specific to this study...................................................................................................... 111

Figure 37. Number of responses as a function of voter confidence level in cast votes
being counted in the 2012 Presidential Election totals ........................................... 116

Figure 38. Number of responses as a function of likelihood that voters would check-on
their votes through an online verification system.................................................... 117

Figure 39. Number of responses as a function of preferred form of vote verification. . 118
LIST OF TABLES

Table 1. The number and percent of ballots with one or more errors as a function of voting system type ................................................................. 45

Table 2. Errors and Likely Contributing Design Deficiencies that Led to Helios Vote Casting and Vote Verification Failures.......................................................... 63

Table 3. Errors and Likely Contributing Design Deficiencies that Led to PAV Vote Casting and Vote Verification Failures......................................................... 69

Table 4. Errors and Likely Contributing Design Deficiencies that Led to Scantegrity II Vote Casting and Vote Verification Failures............................................... 74

Table 5. Research Questions (and Their Response Sets) that Are of Central Concern to This Project ........................................................................................................... 114
# LIST OF ABBREVIATIONS

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>DRE</td>
<td>Direct-recording Electronic Voting Machine</td>
</tr>
<tr>
<td>e2e</td>
<td>End-to-end PAV</td>
</tr>
<tr>
<td>GUI</td>
<td>Graphical User Interface</td>
</tr>
<tr>
<td>HCI</td>
<td>Human Computer Interaction</td>
</tr>
<tr>
<td>HF</td>
<td>Human Factors</td>
</tr>
<tr>
<td>PaV</td>
<td>Prêt à Voter</td>
</tr>
<tr>
<td>PIN</td>
<td>Personal Identification Number</td>
</tr>
<tr>
<td>SBT</td>
<td>Smart Ballot Tracker</td>
</tr>
<tr>
<td>SUS</td>
<td>System Usability Scale</td>
</tr>
<tr>
<td>UI</td>
<td>User Interface</td>
</tr>
</tbody>
</table>
CHAPTER 1

Introduction

For centuries there has been a desire for auditability in elections. In mid-19th century America, groups of voters stood in public venues and called out their ballot choices to the election clerks, while a judge tallied the votes (Jones, 2001). The advantage of this voting method was that anyone could listen to the vocal expression of preferences and keep their own vote count, which prevented practices like ballot box stuffing. While this oral voting method may have increased the accuracy of vote counting, voters’ desire for privacy was not addressed, enabling bribery and coercion. In response, during the late 1800s, voting jurisdictions began to introduce the use of the secret, Australian ballots that listed all the candidates for the same office on the same sheet of paper (which was issued to voters at the polling station) and guaranteed voters privacy in preparing ballots inside a booth (Brent, 2006). This voting system ensured that voters prepared their own ballot expressing their intent while preserving anonymity. Yet this voting method was not perfect; there was not a means to audit the election—leaving a long-standing tension between auditability and privacy in elections.

e2e Voting Systems

So that cast ballots can be both auditable and anonymous, which would ultimately improve the integrity of elections, voting security researchers have developed secure,
voter verifiable systems, also known as end-to-end (e2e) voting systems (e.g., Adida, 2008; Carback et al., 2010; Chaum et al., 2010; Clarkson, 2008; Ryan et al., 2009). e2e systems are voting methods that aim for ballots to be cast as voters intend and counted as cast. To make sure these systems are functioning as they should, they are designed so that both voters and observers can audit, or verify, various aspects of the voting method—all while preserving voter privacy.

How do these e2e systems work? To protect votes from malicious attacks like remotely changing votes in real-time before they are tallied and altering software to rig an election during the system build process (Constantine, 2014), cryptographic protocols and auditing mechanisms are used. The cryptographic methods make it very difficult to undetectably attack and/or alter the e2e systems so that election outcomes would be impacted. Then, with the ability for voters and observers to audit the systems, people are given a means to make sure the systems are working as they should. To give an example, at a polling station voters can randomly check that an e2e ballot is constructed properly by making selections on a ballot, encrypting it, and then instead of casting it, spoiling it by decrypting it make sure that the original selections are correctly revealed (Chaum et al., 2005). This helps assure people that the system is set up correctly so that their intended selections will be cast. Then to make sure that cast selections are counted accurately, other types of auditing processes may take place. At the polling station voters could be issued a receipt, or they might have to create their own. These receipts never list actual candidates or propositions that the voters selected in order to keep the ballots secret. At the same time, other information is shown on the receipt such as a special code or tracker associated with the ballot, unique codes linked to each selection (and only a
voter knows what the code signifies), or an image of a ballot with marked selections in which the candidates names have been removed. These type of receipts were designed to offer voters a means to prove that they cast a ballot that should be counted in the election totals. Voters can also use the receipt, if they choose, to later check through an online election system that their ballot was recorded by the system. This process allows the voters to make sure that the ballot selections were correctly transmitted to the election bulletin board where they will be tallied.

A non-voter, or any interested third party, also can audit an election that uses e2e systems. For instance, even if they do not cast a vote, in principal anyone can make selections on a ballot, create a receipt, and then decrypt the information to see if it matches the original selections. Then to make sure that the final tally is correct, there is often an intricate process to decrypt votes (but not ballots) across a chain of individuals and/or systems so that the votes can be added up to determine the winner of the race, without ever linking a particular vote to an individual. An interested party may be able to perform a set of operations to make sure the final vote count is accurate.

As can be seen by the instances outlined above, auditing practices that take place randomly throughout the election can help to realize a fair, accurate election in which any person or organization has the ability to audit the system in order to detect errors. In principle, this is a significant improvement over paper and DRE voting methods that do not offer a means for a voter to track their ballot after casting it or for people to make sure the votes are added correctly. Instead, the voter has to trust that their votes have been received and will be processed by the system, and individuals must hope that election authorities will add-up every single ballot correctly. In contrast, when voting with an e2e
voting system, no individuals have to be trusted. This is important because voters, poll workers, election officials, voting system programmers, technicians, and administrators all have the potential to have an interest in achieving specific election outcomes (Jones, 2004). What is more, the same level of privacy is assured with e2e systems as the paper ballot and DRE methods. By always keeping voters’ selections anonymous, practices like vote buying and coercion are inhibited. Therefore, advantages of DRE and paper ballots are retained but are improved upon by giving people a means to track votes and corresponding tallies.

Numerous e2e auditable voting systems have been proposed over the last decade: Secret-ballot Receipts (Chaum, 2004), Scratch & Vote (Adida & Rivest, 2006), ThreeBallot (Rivest, 2006), Bingo Voting (Bohli et al., 2007), Scantegrity II (Chaum et al., 2008), Punchscan (Essex et al, 2008), Helios (Adida, 2008), Prêt à Voter (Ryan et al., 2009), Wombat (Grundland, 2012; Farhi, 2013), Remotegrity (Zagorski et al., 2013), STAR-Vote (Bell et al., 2013), and DEMOS (Delis et al., 2014)—to give some examples. Out of all the e2e systems described in published papers, three systems—Helios, Prêt à Voter, and Scantegrity II—have been selected to be representative examples of voter verifiable systems for several reasons. First, they are largely accepted and discussed as secure voting methods within the voting research community. Furthermore, they represent a spectrum of the different solution types that have been proposed for use in polling stations (Helios can potentially be modified and adapted for use at polling sites in order to prevent coercion). Helios is a web-based system and an exemplar of Benaloh-style schemes (Benaloh, 2006). Prêt à Voter (PaV) is a simple, novel, paper-based scheme with many variants that are being considered for use in various elections all over
the world. Scantegrity II is another paper-based scheme that incorporates the traditional paper bubble ballot. All three voting systems have been used, or will be used, in actual elections: Helios was used in the presidential election at the Universite Catholique de Louvain, Belgium (Adida et al., 2009), International Association for Cryptologic Research’s board of directors election (IACR, n.d.), and Princeton Undergraduate Elections (see princeton.heliosvoting.org). PaV has been used in student elections in both Luxembourg and Surrey (P. Ryan, personal communication, April 3, 2014), and it will be used in the November 2014 Victorian State elections (Burton et al., 2012). Scantegrity II was used in the November 2009 municipal election in Takoma Park, Maryland (Carback et al., 2010).

Helios

Helios is a web-based, open-audit voting system (Adida, 2008; Adida et al., 2009) utilizing peer-reviewed cryptographic techniques that allow voters to cast a ballot with a secure system, with anyone being able to audit the entire process. From a security standpoint, system highlights include browser-based encryption (i.e., ballots are encrypted within the voter’s web browser), homomorphic tallying (i.e., voters’ selections are kept private by never decrypting individual ballots, instead the information is extracted through a distributed process or chain of trustees so that the final tally can be decrypted and checked-on by anyone), user authentication by email address (in order to verify a person is able to vote in an election and make sure that only their last cast ballot
will be counted in the final tally), and vote casting assurance through various levels of auditing as discussed in the previous paragraphs.

From the voter’s standpoint, Helios appears to be similar to direct recording electronic voting systems (DREs) like VoteBox (Sandler, et al, 2008).

Vote Casting. The following outlines the vote casting process from the voter’s perspective (the exact steps have the potential to vary from voter to voter, hence the following are potential procedures):

1. The voter logs into their email account to obtain the election’s website address (this information can also be disseminated through other methods).

2. After navigating to the election’s Helios Voting Booth webpage, the voter reads through the voting system instructions and clicks “start” to begin voting. See Figure 1.

![Figure 1. Screenshot of Helios Voting Booth instructions](image)
3. The voter completes the ballot one race at a time by checking the box next to the desired candidate or proposition and then clicking next/proceed to move onto the next screen. See Figure 2.

![Figure 2. Screenshot of the presidential race on the Helios ballot](image)

4. The voter reviews his or her ballot and then clicks the “confirm choices and encrypt ballot” button. See Figure 3. Encrypting the ballot ensures that other parties or systems cannot see the voter’s actual selections.
5. The voter records his or her smart ballot tracker (SBT) by printing it out and proceeds to submission. The advantage of recording the SBT is that it serves as a link to the ballot that can later be tracked if a voter chooses.

6. The voter logs in with their email address to check their eligibility to vote. See Figure 4. During this step, the system also notes if a previously cast ballot needs to be removed from the final tally since only the last cast ballot is counted.
7. The voter casts the ballot associated with their smart ballot tracker.

8. The voter views a screen indicating their vote has been successfully cast. See Figure 5.

Figure 4. Screenshot of the page that tells voters they must log in to verify their eligibility

Figure 5. Screenshot of the Helios cast vote confirmation page
**Vote Verifying.** For a voter to verify their vote, or check that it was in fact cast in the election, the following sequence is typical:

1. In the user’s inbox, the voter opens and views an email from the Helios Voting Administrator. The email indicates that their vote has been successfully cast and displays a link where the ballot is archived.

2. The voter clicks on the ballot archive link.

3. The voter views a screen that says “Cast Vote” along with their SBT. The voter clicks on details and views the code associated with the ballot, which can be used on an auditing page to verify that their ballot is encrypted correctly. See Figure 6.

![](image)

**Figure 6.** Screenshot of the voter’s archived ballot

4. The voter returns to the election home page and clicks on “Votes and Ballots.”

5. The voter observes on the Voter and Ballot Tracking Center page that their SBT is shown within the list of cast votes. See Figure 7. When a voter sees
their SBT on this page, they are supposed to be assured that the system recorded their votes as cast.

Figure 7. Screenshot of Helios’ Voters and Ballot Tracking Center

Prêt à Voter

The next system, Prêt à Voter (PaV), was inspired by Chaum’s (2004) visual cryptographic scheme, which featured a receipt made up of layers of paper that when viewed individually did not show any meaningful information, but when the layers were superimposed on top of each other, the voter’s selections were revealed. PaV is a voting system that allows voters to vote with paper forms (with randomly ordered races and selections for each race), which can be physically modified to then serve as an encrypted ballot, which keeps the ballot secret. This voting method is auditable at numerous phases.
by both voters and teams of auditors to illustrate at every step of the election whether or not there are any problems (Ryan et al., 2009). The system is flexible in that it allows different encryption schemes and cryptographic mechanisms to be used as needed to meet the requirements of different elections.

PaV was intended to provide voters with a simple, familiar voter experience of completing a paper ballot by placing a mark next to desired candidates and propositions.

*Vote Casting.* To vote with the PaV system, the voter follows these typical steps:

1. A sealed envelope enclosing a paper ballot is given to the voter. The voter opens the envelope and finds an instruction sheet and cards that make up the ballot. See Figure 8 for the instruction sheet.

![General Election Ballot](image)

**Figure 8.** Voting instructions for PaV.
2. To mark their selections on the ballot cards (see Figure 9), a cross (x) is marked in the right hand box next to the name of the candidate or proposition that the voter wants to select.

Figure 9. Card 1/8 of the PaV study ballot
3. After completing the ballot, the voter detaches the candidates lists from their selections or marks. The detachment of the candidate names from the voter’s crosses essentially encrypts the ballots so that a person who sees the selections would not be able to determine the candidates for whom votes were cast.

4. The candidates lists are shredded. In this step the voter destroys evidence of the specific order of his or her randomly ordered candidates list so that the ballot cards cannot be reconstructed.

5. The voter walks over to the vote casting station and feeds the voting slips into the scanner. See Figure 10. By scanning in the ballots, the votes are effectively cast as the system receives, processes, and transmits them to the election bulletin board.

Figure 10. Photograph of a voter casting their PaV ballot
6. The voting slips are placed in the ballot box. This serves as a redundant paper trail in case the original ballots ever need to be referenced.

7. The voter takes their printed receipt, proving that they cast a ballot. The receipt (shown in Figure 11) shows images of the scanned voting slips along with the website and ballot verification code—information needed later if they want to verify their votes.

![General Election Ballot](image)

**Figure 11.** An example of a PaV voter receipt
Vote Verification. For a voter to verify their vote using PaV, the voter might typically perform the following sequence on a computer or mobile device:

1. The voter navigates to the election verification website, which is printed on their receipt.

2. The voter enters the ballot verification code on the home page and submits it. See Figure 12. The verification code is unique to the voter’s particular ballot and therefore acts as a link to the cast ballot cards.

![Texas General Election, November 8, 2016](image)

**Figure 12.** Screenshot of study’s PaV’s vote verification home webpage

3. The voter views the vote validation page confirming the entered verification code is valid. This page also displays images of every ballot card—thereby displaying every selection on every card (without any candidates lists) that makes up their ballot. See Figure 13. This step in the verification process is intended to assure a voter that their particular ballot was transmitted to the system and included in the tally displayed on the election bulletin board.
Figure 13. Screenshot of PaV’s vote validation webpage developed for the research study

Scantegrity II

The third method, Scantegrity II, is an optical scan voting system that enables a voter to vote with a paper bubble ballot, which is enhanced by unique, traceable confirmation codes that can be revealed by invisible ink decoder pens (Chaum et al., 2008). This voting system can be audited by voters or any other interested party.

Scantegrity II was developed so that voters could still use a familiar voting technology: an optical scan bubble ballot that they already have experience using.

Vote Casting. To cast a vote using the Scantegrity II voting method, a voter would typically do the following:
1. The voter reads the instructions on both the ballot (shown in Figure 14) and separate vote verification sheet (shown in Figure 15).

Figure 14. Front of the Scantegrity II ballot
INSTRUCTIONS FOR VERIFYING YOUR VOTE ON-LINE AFTER YOU RETURN HOME

You have the OPTION of verifying your vote on-line after you return home. It is not necessary to do so. You may ignore this step entirely; your cast ballot will be counted whether or not you do this verification process.

If you wish to verify your vote on-line, perform the following steps:

1. Fill out your ballot according to the instructions provided on the ballot. “Confirmation numbers” will appear inside the oval you mark.

2. BEFORE YOU CAST YOUR BALLOT record the Online Verification Number and the confirmation numbers below, using the special pen.

*“On-Line Verification Number” from the bottom right corner of your ballot:

<table>
<thead>
<tr>
<th>Race</th>
<th>Code</th>
<th>Race</th>
<th>Code</th>
</tr>
</thead>
<tbody>
<tr>
<td>President And Vice President</td>
<td></td>
<td>Judge Texas Supreme Court</td>
<td></td>
</tr>
<tr>
<td>United States Senator</td>
<td></td>
<td>Judge Court of Criminal Appeals</td>
<td></td>
</tr>
<tr>
<td>Representative in Congress</td>
<td></td>
<td>District Attorney</td>
<td></td>
</tr>
<tr>
<td>Governor</td>
<td></td>
<td>County Treasurer</td>
<td></td>
</tr>
<tr>
<td>Lieutenant Governor</td>
<td></td>
<td>Sheriff</td>
<td></td>
</tr>
<tr>
<td>Attorney General</td>
<td></td>
<td>County Tax Assessor</td>
<td></td>
</tr>
<tr>
<td>Comptroller of Public Accounts</td>
<td></td>
<td>Justice of the Peace</td>
<td></td>
</tr>
<tr>
<td>Commissioner of General Land Office</td>
<td></td>
<td>County Judge</td>
<td>Proposition 1</td>
</tr>
<tr>
<td>Commissioner of Agriculture</td>
<td></td>
<td>Proposition 2</td>
<td></td>
</tr>
<tr>
<td>Railroad Commissioner</td>
<td></td>
<td>Proposition 3</td>
<td></td>
</tr>
<tr>
<td>State Senator</td>
<td></td>
<td>Proposition 4</td>
<td></td>
</tr>
<tr>
<td>State Representative District 134</td>
<td></td>
<td>Proposition 5</td>
<td></td>
</tr>
<tr>
<td>Member State Board of Education,</td>
<td></td>
<td>Proposition 6</td>
<td></td>
</tr>
<tr>
<td>District 2</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

3. Cast your ballot as usual using the polling station’s scanner. DO NOT CAST THIS SHEET, but take it home with you.

4. After you have returned home, use a computer with an Internet connection to access the County’s vote verification web page: mockelection.rice.edu. Here you will see instructions for verifying that the confirmation numbers you wrote down are correctly recorded. Note that the confirmation numbers are randomly generated and cannot be used to determine how you voted.

Figure 15. Scantegrity II vote verification sheet

2. The voter uses the special marking device to make ballot selections—and consequently reveal codes—by filling in the appropriate bubbles. See Figure 16 for an example of a completed ballot. While using decoder pens to reveal information printed with invisible ink may be good “security theater,” this feature in itself does not actually enhance the system’s
security. In contrast, the codes associated with each selection are designed to give the voter a means to track their votes throughout the election while still keeping them anonymous.

![Completed Scantegrity II Ballot](image)

**Figure 16.** Photograph of a completed Scantegrity II ballot, with invisible ink confirmation codes revealed.
3. The voter records on the separate vote verification sheet the revealed confirmation codes found inside each marked bubble. He or she also records on this sheet the ballot ID / online verification number that is found on the bottom right corner of the ballot. In this step, voters are creating an encrypted receipt that gives them a link to their particular ballot and a means to record each selection they made (via the code). The idea is that the codes will be meaningful to the voter but no one else.

4. The voter walks over to the ballot casting station to scan in the ballot and have it then placed in the ballot box. Like PaV, the voting system records the ballot and tallies the results, removing the chance of human error or intentional alterations to the vote count.

5. The voter hands the vote verification sheet they completed to the polling station official so that they can stamp “Cast Ballot” on it. Marking the receipt with the stamp signifies to the voter, or others, that the receipt corresponds to a cast ballot, versus a spoiled ballot.

6. The voter chooses whether or not to keep their verification sheet, depending on whether or not they might want to check on their ballot at a later time. As with all three e2e systems, the verification step is entirely optional, so some voters may not want to keep their receipt.

*Vote Verification.* To verify the votes, a voter may perform the following sequence at their home or office:
1. The voter navigates to the election’s vote verification web page shown in Figure 17. (This verification system allows anyone to check on the election results.)

![Harris County General Election, November 8, 2016](image)

**Figure 17.** Screenshot of Scantegrity II’s vote verification home page developed for this study

2. The voter enters their unique online verification number associated with their ballot.

3. The voter views a confirmation webpage that says the ballot has been cast and processed. This page also displays the online validation code along with a list of the voter’s confirmation codes, with each code corresponding to a ballot selection. See Figure 18. This is the opportunity for the voter to make sure that the system recorded all of their votes and posted them correctly. If a voter notices some of the displayed codes do not match those transcribed onto their receipt, then they should report this finding to election officials.
As can be seen, all three of these e2e, verifiable systems are theoretically an improvement over non-e2e voting methods. Through evidence the system provides, voters and others who may be interested can observe that the system is accurate, reliable, and secure.

**Research Project Description**

Despite voting security experts’ best intentions to improve the voting systems, it is not yet well understood how voters’ performance, behavior, and impressions might be impacted after their implementation. For example, voters already have difficulty voting
with standard paper ballots (Norden, Kimball, Quesenbery & Chen, 2008) due to inadequate instructions and confusing ballot design. If additional e2e mechanisms are then laid on top of these problems, one must question if voters’ abilities to cast their vote will be further degraded. At present there is not a clear understanding of whether voters will be able to understand why these new procedures have been added to the voting process, and if the secure methods really do impact perceptions of these systems. Not knowing this type of information about e2e systems is problematic. If the addition of these enhanced security mechanisms makes voting too hard, or too complex to understand, then voters might be discouraged from voting, become disenfranchised, and fail to participate in future elections—ultimately impacting the outcome of races.

There have been a few studies that have attempted to study usability (i.e., how easy or hard the systems were to use) of select e2e systems. Of those that have, one study by Weber & Hengartner (2009) aimed to study the usability of Helios. The researchers held a mock undergraduate student election in which 20 participants voted with a ballot composed of seven races. While the participants voted, they were asked to describe out loud what they were thinking (i.e., perform a think out loud protocol). The researchers did not record specific performance measures. Rather they recorded the researchers observations, the voters’ verbal comments, and participant responses from the post-use interview. This data was used to identify “themes” that kept arising during voting, auditing, and verifying processes. Examples of these themes included user struggles (e.g., voters loosing the completed Helios ballot due to timeout errors while the system was encrypting the ballot, and the system acting differently than the users expected when they tried to audit their ballot), perceived benefits of the system (e.g., convenience, accuracy,
and efficiency), and if the voters thought the system was secure and accurate (i.e., most participants were confident that their votes would be kept secret and were accurately recorded). Besides trying to study the usability of Helios from the voter’s point of view, the researchers also stood in as the election administrators and analyzed the usability of the system from their perspective (i.e., cognitive walkthrough method). Based on these types of collected data, the researchers made usability assessments. The researchers found about half of the participants did not cast a vote, there were discrepancies between the final vote tally generated by the Helios voting system and the count arrived at by the system users, voters seemed to think the system was accurate, and the participants did not fully understand the language used to describe system security. In general, these findings suggested that the Helios voting system was not fully usable. Consequently, the researchers made some recommendations for system improvements like changing the interface so that radio buttons are used instead of checkboxes when completing a ballot and reducing the number of script time-out warnings during the encryption process.

A second study by Karayumak and colleagues (2011) also aimed to study the usability of Helios. In this study a group of security, electronic voting, and usability experts performed a cognitive walkthrough over several sessions to assess the vote casting and vote verification methods. Particularly, they first identified usability problems such as wording that was too technical for voters to understand, changing the way the voter interacts with the interface throughout the voting process, providing a “help” link that only gives the user a means to send an email versus providing immediate assistance, and the presentation of complex information that was not necessary to cast a vote. Next, the researchers mapped each usability problem to the seven principles of design outlined
by ISO 9241-210: suitability for the task, self-descriptiveness, conformity with user expectations, suitability for learning, controllability, error tolerance, and suitability for individualization. Last, the researchers made concrete suggestions as to how to improve the interface. Some of their suggested improvements included simplifying voter instructions, shortening the verification code (SBT), and redesigning the user interface so it is consistent from page to page.

In a third usability study, Winckler and colleagues (2009) ran a field trial at Newcastle University to study the usability of PaV. 105 people voted with the system in an election for which the winning charity would receive a donation. Usability was described as the “perceived ease of use and usefulness of [an] information technology system” (p. 5). Their measures of usability included the System Usability Scale (SUS), along with “5 questions directly related to the voting procedure (e.g. It was clear to me why I need to destroy half of the voting ballot” (p. 7). A review of the data revealed that a mean SUS score for PaV was 68.5, with a standard deviation of 17.8 (only 65 of the participants completed the instrument). 75.8% of the participants indicated that they fully understood the voting instructions. Roughly half of the participants felt they understood why a portion of the ballot had to be destroyed, that the receipt corresponded with their marked ballot, that the voting method was trustworthy, and that they felt confident using the voting method. The authors of the paper recognized in their conclusion that there was room to improve the usability of PaV is future iterations, especially since the SUS score from this study was below other voting systems that researchers like Byrne et al. (2007) tested in prior research.
In a fourth study by Carback and colleagues (2010), voters’ experiences using Scantegrity II in the November 2009 local election in Takoma Park were assessed. The system users were actual voters, but the residents were not fully representative of the U.S. population because the voters were primarily affluent and highly educated. The authors expressed the work was not meant to be a usability study. Nonetheless, the researchers recorded whether voters could cast votes, how long it took them to vote, observed errors and failures, the number of participants who verified their vote, and asked questions about ease of system use and confidence in parts of the system. The researchers found that the voters and election officials could use Scantegrity II. However, many voters had to ask for assistance since they did not know how to complete a ballot or create a receipt, the double ended pens confused them, they had trouble operating the scanner, and/or felt they already knew how to vote and hence refused to read the instructions for the new voting method and then proceeded to make mistakes. If the poll workers had not provided assistance to the voters, it is likely that vote casting rates would have been impacted. As for time on task to vote, 93 or 5% of voters were timed from the moment they were handed a ballot to the moment they walked away from the scanners. The mean voting time was 167 seconds, with a range from 55 seconds to 10 minutes. It should be noted that the ballot only had races for mayor and city council positions. It is also not known how many of the participants in this study recorded all of the codes and ballot information required to create a self-generated receipt. If a large portion of voters did not do this, voting times could end up being longer. As for the voters who verified before the deadline to review ballots and file complaints, 66 or 4% of the total ballots were verified online. This number was determined to be sufficient to detect errors and fraud. As for the
voter experiences, the majority of the participants strongly agreed that the system was both easy to use and that they had confidence in the system. These findings seemed to contradict the poll workers who pointed out that there was “too much information,” some voters did not know what to do, and the equipment was hard to use or confusing. Despite problems encountered while using the system, the researchers concluded the election to be a success since a secure, verifiable voting system that keeps voters’ selections private was used in an actual election.

While the above studies have reported on the usability of e2e voting systems, none have evaluated the voting methods using a sample of diverse voters and longer ballots that are representative of those used to vote in a U.S. presidential election. Prior studies also have not methodically measured all three dimensions of usability described by ISO 9241-1: effectiveness, efficiency and satisfaction (ISO, 1998). The benefit of using ISO 9241-11 is that it is the gold standard for measuring usability in an experimental research setting. Furthermore, it allows for e2e voting systems to be compared to previous research in this area. The literature review also highlights that few researchers have evaluated voters’ impressions of these systems, or assessed if voters understand why they must complete the novel procedures required vote and verify using the e2e methods. Even when perceptions and understanding have been studied, the findings have not been compared across voting systems. Last, a review of the previous studies illustrated that research has not been conducted to assess if voters would actually want to use e2e systems in a large-scale election.

To better understand these issues, this research project evaluated the usability of the three representative e2e voting schemes—Helios, PsV, and Scantegrity II—per ISO
9241-11 criteria, identified potential contributing design deficiencies linked to the observed usability failures, determined if voters understand how each procedure makes the voting method secure, and assessed voters’ impressions and anticipated use of each system. The participants were representative of a diverse voting population, and the ballots were made up of 27 races and propositions, which is more representative of ballots used in U.S. government elections. The following describes in more detail the three studies completed to answer specific research questions regarding e2e system use and associated mental models:

**Study 1—Part 1 of 2**

Part one of the first study asked six questions to understand the usability of the three e2e systems: 1) Can voters who use the e2e systems to mark their ballots as they intend? 2) Are voters able to cast their completed ballot? 3) Do these e2e systems significantly impact the ability of voters to vote in a timely fashion? 4) Are voters able to verify their votes after casting their ballot? 5) How long does it take voters to complete the vote verification process? Last, 6) What are voters’ subjective usability assessments of these more complicated, but verifiable, voting systems? In the process of answering these questions, baseline usability data were collected. This data can be used to both compare these e2e systems to voting technologies previously tested and determine if voters can easily use the systems. If the systems perform below other voting technologies and/or voters struggle to use them, then researchers are given the opportunity to improve upon them before they are used in large-scale, government elections.
Study 1—Part 2 of 2

In the second part of the first study, two pointed questions were asked: 1) If voters failed at vote casting and/or vote verifying, why were they not able to complete these two basic tasks? And 2) what were the likely contributing system design deficiencies? These questions were asked to begin identifying concrete causes of usability problems. By focusing on the failures, as opposed to every user struggle encountered in the vote casting and vote verifying processes, it is hoped that the systems can be improved enough in the next generations so that voters can at least complete their most basic goals.

Study 2

Three questions were answered in the second study: 1) Do users understand how the secure systems work? 2) What kinds of mental models do voters form regarding the processes that lead to enhanced security? And 3) how do these mental models differ from those voters have for more typical voting technologies like the paper ballot? Answering these types of questions can lead to a more satisfactory user experience. By having a basic understanding of how voters think about e2e voting methods, system designers can then make sure they design the systems for how users actually think about them, as opposed to how the designers believe the voters think about them. Moreover, if it is found that voters are frustrated by the novel, complicated e2e system features, and the system users do not understand why they are executing certain procedures, then these identified issues can be addressed in future system iterations.
Study 3

In the third study, three research questions addressed voters’ interest in using a verification system after voting. It was asked 1) how confident are voters in their votes being cast and counted accurately with current voting technologies? 2) Would voters even want to use post-voting features to check on their votes? And 3) what form of verification would voters like to see when checking on their ballot? Even though anticipated system use may differ from actual system use, if users do not even want to use a system to fulfill a goal, then the system might need to be reconsidered. In the context of e2e voting systems, the developers always include a means for voters to check on their ballots after they are cast. This study is motivated by the fact that it is unknown if voters actually want to do this and the type of information that they would want to help them with the process.

In this document, each of these studies (and study parts) is presented in its own chapter, written to be able to stand on its own. For detailed information about background information, concepts studied, methods, results, discussions, and conclusions related to each study, the reader should refer to the individual chapters. The global discussion at the end of this dissertation provides a general overview of the research findings and directions to move forward with future research and development of e2e systems.

The information about e2e voting systems acquired through this research project will begin to fill gaps in knowledge regarding users’ interactions with these voting methods, and their understanding of them. Once researchers better understand how voters generally use, perceive, and think about e2e systems, design recommendations can be
made. These recommendations will specify how to improve upon and develop these types of systems in the future, which ultimately will improve the integrity of elections.

This research project is timely because all of the systems are in early stages of development and have not been implemented in large-scale national or state elections. Therefore, system usability and mental model findings would be valuable to system designers and researchers as they seek to improve and implement secure voting systems that are usable by every voter—despite differences across voters like age, education, socio-economic background, disability, or reading ability.
CHAPTER 2

Usability of Voter Verifiable End-to-end Voting Systems:
Baseline Data for Helios, Prêt à Voter, and Scantegrity II

INTRODUCTION

Did the auditing solutions to voting security and accuracy problems lead to an entirely new set of problems? This might be the case, and researchers need to find out. Voting systems like the paper ballot and electronic voting machine (or direct-recording electronic voting system, DRE) are not secure and auditable. Hence voting security researchers attempted to solve these issues by developing voter-verifiable end-to-end (e2e) systems—as described in the previous chapter—such as Helios, Prêt à Voter (PAV), and Scantegrity II (e.g., Adida, 2008; Ryan et al., 2009; Chaum et al., 2008). In the process of designing these systems to be resistant to malicious attacks, auditable by any interested party, and capable of maintaining voter privacy, the systems seem to have become unusual and complex from voters’ perspectives when compared to previous voting technologies.

To vote with Helios requires a voter to complete a multistep encryption process, verify through an email address that they are eligible to vote in an election, and record a long ballot tracker so that they can check at an online bulletin board that their vote has been captured by the system. PaV requires voters to essentially tear their ballots in half
(i.e. detach their selections from the candidates lists), shred the correct half (i.e., the
candidates list), scan in their ballot, wait for and then keep a system generated receipt that
shows images of their actual ballot cards, and then compare the images on their receipt to
what is shown on the election’s website. With Scantegrity II, a voter uses decoder pens to
reveal invisible ink codes, completes a special vote confirmation sheet by transcribing a
unique ballot ID and revealed codes found inside each selected bubble, inserts their ballot
into a scanner to cast it, and then uses an online verification system to compare the
information on their self-generated receipt to the display shown on the webpage that is
associated with their ballot ID. Absent from these descriptions are the actual ballot
completion steps, which is essentially all a voter had to do with a paper ballot before
dropping it into a locked ballot box. If these systems are as confusing and cumbersome to
voters as they appear that they might be, it is easy to see how a new set of system
usability problems might need to be addressed.

**Understanding the Usability of e2e Voting Systems**

As can be seen from the vote casting and vote verification procedures briefly
described above and discussed in Chapter 1, the three e2e systems are generally complex
from the standpoint of the voter. Many of the processes required to use the systems are
both long and novel in the context of voting. This is of concern because voters cannot
participate in an election, they will become disenfranchised and election outcomes might
be changed—tremendous threats to democracy. Furthermore, if people are not able to
verify that their ballot has been cast because the system is too hard to use, then the system
is not auditable—leaving room for inaccuracy and corruption. Consequently, voting
researchers need to understand the usability of each system and how it compares to other voting technologies in order to make recommendations and impact policy.

System usability is defined as the capability of a range of users to be able to easily and effectively fulfill a specified range of tasks within specified environmental scenarios (Shackel, 1991). In the context of voting, usability might be thought of as whether or not voters can use a voting method to successfully cast their votes. Per ISO standard 9241-11 (1998), there are three suggested measurements of usability: effectiveness, efficiency and satisfaction. As established in previous voting usability research (Byrne et al., 2007; Laskowski et al., 2004), effectiveness addresses whether or not voters are able to select, without error, the candidate or proposition for which they intend to vote. One way to measure effectiveness is by calculating error rates. Efficiency concerns the amount of resources required of a voter to attempt achieving his or her goal. This variable can be measured by calculating task completion times, or the amount of time it takes to vote or verify a vote. The third measure, satisfaction, is defined as the voter’s subjective perceptions of a voting system after using it—such as how hard or easy it is to vote using the method. Satisfaction can be measured with a standardized instrument like the System Usability Scale, or SUS (Brooke, 1996).

The only way to know if e2e systems are usable is to empirically test them. While other studies have reported on the usability of select e2e systems (Carback et al., 2010; Karayumak et al., 2011; Weber et al., 2009; Winckler et al., 2009), none have experimentally evaluated the voting methods along all three suggested measurements outlined by both ISO standard 9241-11 and the 2004 NIST report on voting system usability (Laskowski et al., 2004). In addition, the previous studies neither used diverse
samples that are truly representative of U.S. voters nor longer ballots like those typically used in U.S. elections.

To address this lacuna, this study tested the usability of the three representative e2e voting systems presented above: Helios, Prêt à Voter, and Scantegrity II. When applicable, the same materials and protocols were used from the previous voting studies conducted by Rice University’s human factors voting laboratories (e.g., Byrne et al., 2007; Campbell et al., 2009; Campbell et al., 2011; Everett, 2007; Everett et al., 2008; Holmes & Kortum, 2013) to allow for comparison of usability findings across different voting technologies. The goals of this research study were to understand whether voters can use these e2e voting methods to cast and verify their votes, identify system attributes that might be preventing voters from fulfilling their goals of vote casting and verifying, and help to make recommendations that might enhance the design and implementation of e2e systems.

METHODS

Participants

Thirty-seven participants who were U.S. citizens and 18 years or older (the minimum age to vote in the U.S.) were recruited through an online advertisement in Houston, Texas. They were paid $40 for participating in the study. The mean age was 37.1 years, with a median of 35 and a range of 21 to 64. There were 22 male and 15 female participants. Participants were African American (14, 38%), Caucasian (10, 27%), Mexican American / Chicano (4, 11%), Hispanic / Latino (4, 11%), and other ethnicities
(5, 13%). As for the participants’ educational background, 2 (5%) had completed high school or the GED, 23 (62%) completed some college or an associate’s degree, 8 (22%) were awarded a bachelor’s degree or equivalent, and 4 (11%) held a post-graduate degree. English was the native language of 36 of these participants. All had self-reported normal or corrected-to-normal vision. Participants rated their computer expertise on a scale from 1 to 10, with one being novice and 10 being expert; the mean was 8.2 with a range of 5 to 10. 33 participants had voted in at least one national election, with an average of 3.8 and a range of 0 to 21. Participants had, on average, voted in 5.1 state and local elections. This is a diverse and representative sample of real voters.

Design

A within-subjects design was used, in which every participant used three different voting methods. The within-subjects study design increased the statistical power of the analysis such that the sample size of 37 was more than adequate to detect even small effects. The three voting systems used in this experiment were Helios, Prêt à Voter, and Scantegrity II. Each participant voted with all three methods. All possible orders of presentation were used, and subjects were randomly assigned an order.

So that voters knew for whom they should vote, they were given a list of candidates and propositions. Their list was either primarily Republican and contained 85% Republican candidates, or it was primarily Democratic with 85% being Democratic candidates. Both lists had “yes” votes for four propositions and “no” votes for two. These two lists were the same as those used in previous voting studies conducted at Rice University. Participants were randomly assigned one of the two slates.
Per the ISO 9241-11 definition of usability (ISO, 1998), there were three main dependent variables: errors (effectiveness), completion time (efficiency), and subjective usability (satisfaction). Three types of errors were included in the effectiveness measure. First, I measured the inability to either cast a ballot and/or later verify votes. For example, if a participant completed a ballot but never cast it by scanning it, then this was counted as an error with PaV and Scantegrity II. In Helios, if a voter encrypted his or her ballot but never continued on to verify their eligibility to vote (by logging in with their email account)—an action that is required at this point in the voting process in order to move onto the actual vote casting step, then this would be counted as a failure to cast. Second, per-race errors, which are defined as deviations on the voter’s ballots from the list of candidates and propositions that the voter was instructed to vote for, were recorded. A per-contest error rate for each ballot was computed for every participant. Third, overall ballot errors were measured. Overall ballot errors are defined as a ballot with at least one deviation from the list of candidates and propositions given to the voter. For example, whether a voter selected one wrong candidate or ten wrong candidates, the ballot would be classified as having errors on it.

To measure efficiency, voting and verification completion times were used. Both voting and vote verification times were measured with a stopwatch. The stopwatch was started after the experimenter said the participant could begin, and it was stopped when the participant indicated that they were finished with their task.

The System Usability Scale was used to measure satisfaction. The SUS contains ten subscales. Each subscale is a 5-point Likert scale that measures an aspect of usability.
The ratings for each subscale are combined to yield a single usability score ranging from 0 to 100, with lower scores being associated with lower subjective usability.

Data were also collected on other factors such as technologies used to vote in previous elections, computer experience, perceptions of voting security, and preferred voting technology.

For each e2e system, the dependent measures described above were collected for both the vote casting portion of the system (i.e., the procedures the voter must go through in order to make their selections on a ballot and successfully cast the ballot), as well as the vote verification portion of the system (i.e., the procedures required of the voter to be able to check that their votes were cast and included in the final election tally). The two portions of the system were examined separately since vote verification is an optional procedure not required to cast a ballot and have it be counted. This study did not explore the usability of the optional auditing processes associated with the systems.

**Procedures**

The study began with participants giving their informed consent. They were then read instructions for the experiment. Subjects were instructed to vote on all three ballots according to their list of candidates and propositions. Because verification is neither currently an option in U.S. elections, nor required to cast a vote with e2e systems, voters were specifically told that they would be asked to verify their vote at the end of the voting process, and that they should take whatever steps were necessary to insure that they could perform this verification step. Participants then voted with one of the three voting methods (order was counterbalanced across participants, all orders used), each in
its own room to prevent confusion as to which equipment was associated with each voting system. After voting on a system, the participants immediately completed the System Usability Scale. When completing the instrument, participants were specifically instructed to evaluate the voting system they had just used. Next, participants verified their vote using the same system and completed another SUS, being explicitly instructed to evaluate only the verification system they just used. They then went through this process for the remaining two systems. At the end of the experiment, participants completed a final survey packet that was composed of 49 questions. The survey covered topics like demographics, computer expertise, previous voting experience, security, voting method comparisons, voting method instructions, and vote verification. Last, participants were debriefed, compensated, and thanked for their time.

The modified form of the System Usability Scale as presented in Bangor et al. (2008) was used to assess subjective usability or satisfaction. In this version of the SUS, the word “cumbersome” is replaced with “awkward.” The word “system” was also replaced with the words “voting system” or “voting method,” and “verification system” or “verification method” as appropriate. This particular change was made based on user feedback from the pilot study’s subjects. Altering the SUS in this way has been shown to have no impact on the scale’s reliability (Sauro, 2011).

It should be noted that the participants’ desktops were mirrored to a monitor that only the experimenter could view in another part of the room. Mirroring the monitors was intended to aid the experimenter in observing the participant’s actions in an unobtrusive fashion. Mirrored monitors also allowed the experimenter to score the errors on Helios’ ballot in real time and determine if voters verified their votes across all three systems.
Materials

For all three systems, the following hardware was used: The computers were Dell Optiplex desktops with 17” monitors. The scanners were VuPoint Solution Magic Wands; these scanners were selected because they would automatically feed and scan sheets of paper inserted by the user. The shredders used were Amazon Basics 8 or 12-sheet automatic shredders. The printers used were the HP Deskjet 1000 (Helios) and the HP LaserJet Pro Laser Printer (PaV), both of which are single-function printers. All computers had Windows XP operating systems and Google Chrome version 32 as the default web browser. This web browser was selected because it was compatible with all voting and verification systems tested in this study. The only icons on the computers’ desktops were the hard drive, trashcan, and Google Chrome.

Candidates and propositions on the ballots were those used in Rice University’s previous voting experiments (e.g., Byrne et al., 2007; Everett et al., 2008). The candidates’ names had been randomly generated through online software. The ballot was comprised of 21 races, which included both national and county contests, and six propositions. The length and composition of the ballot was originally designed to reflect the national average number of races. The format and layout of each system’s ballot followed the criteria outlined by the system developers in published papers.

The Helios voting system and election was set up and run through Helios’ website at vote.heliosvoting.org during the winter of 2013-2014. A Gmail login provided to the participant was used to obtain Helios voting instructions, access the election link, confirm eligibility/identity before casting the ballot, and/or view the confirmation email sent after
ballot casting. See Appendix 1 for the study materials used in association with this voting system.

Since PaV had not been previously developed to be used in an election with numerous races (as is the case in the United States), the system was developed based on published papers about PaV (e.g., Lundin & Ryan, 2008; Ryan et al., 2009; Ryan & Peacock, 2010; Ryan & Schneider, 2006), the PaV website (Prêt à Voter, n.d.), and in consultation with Peter Ryan, who first created the system. It should be noted that the security mechanisms were not implemented in the system. Nevertheless, from the voter’s perspective, the system appeared to operate as a fully functional, secure system. See Appendix 2 for system materials.

This study’s implementation of Scantegrity II was heavily based on materials used in the 2009 Takoma Park, Maryland election, in which voters used the system to elect the mayor and city council members (Carback et al., 2010). I also referred to published articles about the system and corresponded through email with Aleks Essex, a researcher who has direct experience with the implementation. When aspects of the system that might have potential to impact usability were not specified, best practices in human factors were followed. Also, when possible, every effort was made to keep system properties (such as font) constant across systems. Like PaV, this system was not a fully functional prototype from a security perspective. Instead, it appeared to be fully functional from the voter’s perspective. See Appendix 3 for Scantegrity II’s materials.
RESULTS

There were no differences in the findings based on whether participants were told to vote for mostly Republicans or mostly Democrats according to their directed voting list. There were no differences in the efficiency, effectiveness, and satisfaction findings based on whether or not participants were able to cast a vote or later verify a vote. In addition, there was no evidence to support system order effects. All of these were treated as single conditions. The analysis was a repeated measures ANOVA unless otherwise specified. $p$-values were adjusted by Greenhouse-Geisser (G-G) correction when appropriate. FDR adjustments to post-hoc tests were performed when necessary.

Vote Casting

Effectiveness

Figure 19 shows the percentage of voters who thought they cast a vote with each system versus the number of actual cast votes. As can be seen, a reliably higher percentage of voters thought they had cast a vote that would be counted in election totals than the percentage of ballots that they actually cast, (tested with binomial linear mixed model, $z = 4.42, p < .001$). The interaction between these two variables across voting systems was not reliable, $z = 0.78, p = .43$. These completion rate findings are extremely troubling. If the tested e2e voting systems are used in a real election, on a large scale, high percentages of voters might not be able to vote—resulting in disastrous outcomes. These failure-to-cast findings are especially unacceptable when many of the other
systems tested in Rice University’s Computer-human Interaction Lab produced 100% ballot casting completion rates (e.g., Byrne et al., 2007).

**Figure 19.** Percentage of cast ballots as a function of voting system, with different colored bars representing perceived and actual cast vote

Per-contest error rates as a function of system can be seen in Figure 20. There was no reliable evidence for an effect of system type on these errors, \( F(1.1, 40.9) = 2.70, MSE = 0.00, p = .10, \eta^2 = .09 \). In this regard, e2e systems seem to be performing better than previously tested voting systems that had error rates ranging from less than 0.5% to about 3.5% (Byrne et al., 2007). With that being said, this potential advantage over other voting technologies is moot if voters cannot cast votes at reasonable rates.

Table 1 shows the frequency of error-containing ballots by voting system. Overall, 5 of the 111 (5%) ballots collected contained at least one error. Again, this error rate is lower than those previously reported (see Byrne et al., 2007). Based on both the
per-contest error rates and error rates by ballot, voters using e2e systems make few errors selecting candidates and propositions on their ballots.

![Mean per-contest error rate percentage as a function of voting system type, with error bars depicting the standard error of the mean.

**Figure 20.** Mean per-contest error rate percentage as a function of voting system type, with error bars depicting the standard error of the mean.

**Table 1.** The number and percent of ballots with one or more errors as a function of voting system type.

<table>
<thead>
<tr>
<th>Voting System</th>
<th>Number of Ballots with Errors</th>
<th>Helios  (3%)</th>
<th>PaV    (11%)</th>
<th>Scantegrity II (0%)</th>
</tr>
</thead>
</table>

**Efficiency**

Average ballot completion time as a function of voting system is presented in Figure 21. As can be seen, there are differences in voting times across the systems, $F(2, 72) = 8.45, MSE = 34,457, p = .001, \eta^2 = .23$. Pairwise tests revealed all three means were
reliably different (Helios and PaV—\(t(36) = 2.14, p = .039\); Helios and Scantegrity II—\(t(36) = 3.47, p = .001\); PaV and Scantegrity II—\(t(36) = 2.48, p = .018\)). Participants took the least amount of time to vote with Helios and the most amount of time to vote with Scantegrity II. In prior research, ballot completion time is generally not sensitive to voting technology. Average completion time for the identical ballot using arrow ballot, bubble ballot, punch card, and lever machine voting methods is approximately 231 seconds (Byrne et al., 2007) and 290 seconds across sequential DRE, direct DRE, bubble ballot, lever machine, and punch card systems (Everett et al., 2008). Thus, the e2e systems impose a substantial time cost on voters.

**Figure 21.** Vote casting completion time as a function of voting system, with dashed horizontal lines depicting the mean and solid horizontal lines depicting the median.
Satisfaction

As can be seen in Figure 22, SUS ratings (out of 100 possible points) differ across the three e2e voting systems, $F(2, 72) = 5.28$, $MSE = 624$, $p = .007$, $\eta^2 = .13$. Pairwise $t$-tests revealed that participants were reliably more satisfied with the usability of Helios (Helios and PaV—$t(36) = 2.91$, $p = .006$; Helios and Scantegrity II—$t(36) = 2.63$, $p = .012$), but there was not a statistically reliable difference in satisfaction ratings between PaV and Scantegrity II ($t(36) = .378$, $p = .71$). When compared to previously tested voting methods, these SUS scores are comparable or lower than those previously seen (Byrne et al., 2007).

**Figure 22.** SUS rating as a function of voting system, with dashed horizontal lines depicting the mean and solid horizontal lines depicting the median
Using the assessment of fitness for use scale (based on the SUS score) proposed by Bangor, Kortum and Miller (2009), Helios would be judged as “acceptable,” while PaV and Scantegrity II would be on the low end of “marginal acceptability.” Based on all of these SUS findings, voters’ satisfaction with using Helios was relatively good, but their satisfaction with using the other two systems was between poor and good—suggesting that there is room for improvement in future system iterations.

**Vote Verification**

**Effectiveness**

Figure 23 shows the number of participants who were able to actually verify their vote through any means versus those who thought they verified as a function of system type. There was no reliable effect of system or difference between perceived versus actual completion rates, $z = 1.66, p = .098$ and $z = 2.17, p = .30$ respectively (again, tested via binomial linear mixed model). However, these vote verification task completion rates are lower than those for vote casting.

With Helios, 16 (43%) voters performed any type of vote verification action. Of these, only 8 (50%) recorded their smart ballot tracker, which allows them to identify their particular vote in the online vote center. Two of the 16 participants verified by viewing the verification email sent to them after voting. The rest of the subjects verified by viewing their information on the Helios election website, keeping in mind that many did not have a recorded smart ballot tracker to which they could refer. With Scantegrity II, 14 (38%) voters performed some type of vote verification. Of these, only nine
attempted to record all 27 vote verification codes; only a single person wrote down all 27 correctly. Based on these results, for both Helios and Scantegrity II participants engaged in a wide range of behaviors when they tried to check that their vote was cast in the mock elections. PaV was designed so that the verification output required to check on the ballot was automatically given to voters upon casting their ballots, and there was only one way in which they could check on their ballots, so more specific findings on verification actions are not reported for the system.

Figure 23. Percentage of verified votes as a function of voting system, with different colored bars representing perceived and actual verified votes

**Efficiency**

Results for vote verification time as a function of voting system are presented in Figure 24. The effect of voting system was suggestive but not statistically reliable, $F(1.2,$
7.2) = 3.74, \( MSE = 21,559, p = .09, \eta^2 = .38 \). It should be noted that the amount of time it takes someone to verify their vote with these e2e voting systems is similar to the amount of time it takes to vote on previously tested voting technologies (Byrne et al., 2007).

**Figure 24.** Verification completion time as a function of voting system, with dashed horizontal lines depicting the mean and solid horizontal lines depicting the median

Satisfaction

Figure 25 depicts the mean SUS score as a function of system type. The effect of voting system was reliable, \( F(2, 12) = 7.86, MSE = 792, p = .007, \eta^2 = .57 \). Pairwise \( t \)-tests indicated that Helios was rated lower than PaV on the subjective usability measure \( (t(6) = 3.50, p = .013) \); there was not any evidence to support other statistically reliable differences (Helios and Scantegrity II: \( t(6) = 3.01, p = .024 \)—this \( t \)-test was not
significant after being FDR adjusted; PaV and Scantegrity II: ($t(6) = 0.43, p = .68$). Using the assessment of fitness for use scale (Bangor et al., 2009), Helios would be judged as being “not acceptable,” Scantegrity II would be on the high end of “marginal,” and PaV would be classified as “good.” To summarize these findings, Helios’ verification system had a staggeringly low subjective usability rating, emphasizing how bad participants thought of the system’s usability. Participants did rate PaV higher. That is, that they thought PaV was easier to use.

**Figure 25.** SUS rating for the vote verification process as a function of voting system, with dashed horizontal lines depicting the mean and solid horizontal lines depicting the median.
DISCUSSION

Generally, all of the tested e2e voting systems appear to have momentous usability issues based just on the high failure-to-cast rates. Perhaps more troubling, however, is the fact that many of the participants in this study thought they cast a vote, but actually did not. These findings would have huge implications in a real election. Since they believe they did in fact vote, they would not even know to tell someone that they could not cast a vote to receive assistance or notify officials that there might be usability problems. As for the voters who recognize they cannot vote, they might seek help or they might give up. Even if they are able to eventually cast a vote after receiving direction, they might choose not to vote in the future, and thus the e2e systems would disenfranchise voters.

The low success rates observed in the vote verification part of the systems are also troublesome. If voters cannot check on their ballot after voting, then fewer people will be able to check that the system is working properly. The voter might also have lower confidence in the system since they know the verification feature is available, but they were not able to use it for some reason. Even if a voter is able to verify that his or her vote was cast, it might lead to frustration levels that are associated with future system avoidance, meaning—again—there will be fewer people to check on the integrity of the system. One potentially unintended consequence of these verification systems is that it adds another opportunity for errors to be committed. If the voters write down their verification information incorrectly (a smart ballot tracker in the case of Helios or a selection’s confirmation code with Scantegrity II) then they might think their vote was
lost, thrown out, or not recorded correctly. If the voter then reports to an election official that something is wrong, a new set of serious problems emerge: election officials and voters might think the election results are incorrect, when in fact they are correct. If widespread, this kind of simple and foreseeable failure could lead to a general lack of confidence in the results among the “average” voter who tried to verify their vote, but failed. These are all serious ramifications—highlighting that it is not enough for a system to be secure. Every system must also be usable.

**Why are these systems failing?**

It is clear that while the e2e mechanisms may significantly enhance the security of these voting systems, the enhancements come at the cost of usability. The additional and unfamiliar procedures impact the very essence of the voting process—the ability to cast a vote—and do so in ways that cause many users to not even be aware that they have failed. I believe that there are several general design choices that led to the results reported here, yet each of these can be overcome with design modifications and additional research efforts.

*1) Security Isn’t Invisible.* All of the tested e2e voting systems function in a way that require users to be an active part of the security process. These additional steps likely lead to increased cognitive load for the user, and that increased load can lead to failures. In contrast, an ideal security mechanism requires no such additional effort on the part of the user. In novice parlance, “it just happens.” The user is neither required to take action nor even know that there is enhanced security
implemented on his behalf. For example, banks encrypt their web-based transactions, but the user does not take part in enabling or executing these additional safety measures.

2) Tested e2e Systems Do Not Model Current Systems to the Greatest Degree Possible. Many of the observed usability difficulties in this study can likely be attributed to designs that work differently than users expect. Many participants were experienced with voting and had seen previous (albeit, different) implementations of what a voting system “should” look like and how it “should” behave. For the most part, the tested e2e systems deviated from these expectations significantly, leaving users confused. In this confusion, participants might have recalled their previous experience with voting systems, and then used that to guide their interactions. Since their previously used voting systems do not work in the same way as e2e voting systems, referring to previous experience inevitably led to decreases in performance and the commission of errors where the users’ prior voting model and the system’s actual function did not match. This may explain why Helios had higher SUS ratings than PaV and Scantegrity II. Many participants verbally expressed that they liked using the computer to vote since they already use them daily—in other words, they got to use a platform with which they were familiar. Of the three systems, Helios also requires the least amount of unfamiliar, novel procedures. Essentially, the voter only has to interact with a series of webpages to vote. In contrast, with PaV voters have to tear their completed ballot in half, shred a portion of it, and then scan what is leftover into a scanner. Scantegrity II is similarly unique, requiring voters to use decoder pens, record revealed invisible ink codes, and then scan in
their ballot. Deviations from the norm can hurt performance and user assessment of that system, which is reflected in the results. Furthermore, PaV and Scantegrity both require that candidate order be randomized, which violates the expectations of most voters and does not conform to election laws in most U.S. jurisdictions.

Even though voters have never seen or interacted with systems like these before, it should not be argued that high rates of failure to cast a vote or to verify a vote are to be expected—hence being acceptable in a system deployed for use. This argument can be countered in two ways. First, completion rates for two previously tested experimental voting systems—IVR and mobile vote—do not suffer from this phenomenon (Holmes & Kortum, 2013; Campbell et al., in press). Second, and more importantly, voting should be considered a walk-up-and-use activity. If a voter only votes in national elections, then there are four years between each interaction a voter has with a particular system, and learning retention is poor under infrequent exposures. Voters must be able to use the system with near 100% success with little or no experience or training.

3) Verification Output Is Not Automated, So Users Make Mistakes. Verification of a vote is a new feature of these systems, so this probably led to some of the system problems like not being able to verify or recognize that their vote had been verified. However, the benefits derived from this feature are so central to these enhanced security systems that more needs to be done to assist voters in the successful completion of this step. As noted, one of the great difficulties users faced is that they either failed to understand that they needed to record additional information to verify, or the additional labor involved dissuaded them from making the effort. Further, even if voters understood and wanted to perform these steps, the likelihood of committing errors in this step was
high. Providing assistance to the voter, such as automated output of the ballot ID (which PaV did) or security codes might have made this step more tenable from the voter’s standpoint.

4) Insufficient User Instructions. Because these e2e system are both relatively new and place additional cognitive burdens on the users, enhanced instruction may be required. This does not necessarily mean giving the voters long, detailed instructions for use at each station, as these were often ignored or skimmed in the systems tested here. It does mean providing specific, clear helping instructions at critical junctures in the process. Instructions should never be a substitute for good design, but occasionally, good inline dialogue can mitigate design features that are crucial to the systems operation. This lack of inline instruction may have been why subjective usability was lowest for Helios. Helios provided instructions in the beginning on how to vote, but after casting a ballot, the system did not tell the voter how they could follow up by verifying to be assured that their vote was handled correctly.

5) Voting Systems Were Not Specified in Detail. One of the things learned quickly as our team tried to construct these systems is that while the security mechanisms were well-specified by the researchers who imagined them, not every system specification was defined. This is understandable, as the papers I used to model e2e systems described the security and general functioning of the system, not every single operational user interface detail. However, anyone (like a county clerk) who wanted to implement such a system would be left to devise their own best practices for all the omitted details, and this could lead to a wide range of outcomes depending on the
implementation. The devil is always in the details, and this is especially true for complex systems such as these. It also points to the need for enhanced collaboration between security researchers and human factors specialists when developing such systems.

**Where do we go from here?**

Despite the usability problems associated with the tested systems, one must keep in mind that they have the potential to be both more secure and more accurate than traditional voting systems once the systems are usable by everyone. Incorporating human factors research and development methods during active system development would be a critical part of ensuring that these types of systems are developed with the user in mind.

There are numerous questions that future research should address. For example, are people with disabilities able to use the voter verifiable systems? If not, what can be done so that they can easily and quickly vote? Are the auditing portions of the system usable? When a voter verifies their vote with a system like Scantegrity II or PaV that displays their unique codes or images of their ballot, how accurate are voters? In other words, would people actually catch errors? How do voters report concerns about their verified votes? All three systems are designed to allow voters to check that things are working properly. But if they are not, what do voters do? By answering questions like these, the systems will be able to be further improved and the relationship between security and usability will be understood in more detail.
CONCLUSION

The data from this study serves as a reference point for future research and discussions about the usability of voter verifiable voting systems. It also enables e2e systems to be compared to other voting systems that have been previously tested or will be tested in the future. With that being said, this study only begins to answer basic research questions surrounding these new systems, while highlighting many avenues for future studies.
CHAPTER 3

From Error to Error: Why Voters Could not Cast a Ballot and Verify Their Vote

INTRODUCTION

Imagine a real, large-scale election in which nearly half of the voters cannot cast a vote. To make matters worse, many of these voters thought they casted a ballot that would be counted in the election tally. As for the voters who realized they could not figure out how to vote, some of them gave up, went home, and wanted to avoid voting in the future. An abysmal scenario like this would not be acceptable as it undermines the democracy of countries by potentially changing election outcomes and disenfranchising voters. While this scenario is made-up, it is not farfetched. The usability study presented in Chapter 2 found that close to half of the participants could not cast a vote using the tested voter verifiable systems in a mock election. Nor were many of the participants able to verify their vote—an optional yet fundamental system feature that helps make sure that the end-to-end systems are accurate and transparent, while keeping individuals’ votes private. Therefore if these tested systems were employed in a real election, it would not be surprising to find similar outcomes.

This dissertation’s usability study illustrates that the Helios, Prêt à Voter (PaV), and Scantegrity II voting systems are difficult to use. However, up to this point, the
specific reasons participants were not able to vote or check on these votes has not been assessed. This second portion of the study aims to identify behavioral errors and system design deficiencies that may have led to these two types of failures. An error is defined in this chapter as the actions, whether intended or not, that prevented the voter from casting a ballot (goal 1) or from checking that their ballot had been cast (goal 2). The system design deficiencies are the system features that contributed to the error occurrence. This chapter does not attempt to address every possible usability struggle, slip, or mistake that might have occurred, or could occur, while vote casting and vote verifying; this list has the potential to be endless (Reason, 1990). For this reason, it should be cautioned that every single possible design deficiency contributing to a failure as outlined in this chapter could be addressed and the systems still might not be usable. (This concept is explored in more detail in this chapter’s discussion section).

It should also be noted that potential, not actual, design deficiencies are being identified in this chapter. Members of the same behavioral error class can arise from quite different causal mechanisms (Reason, 1990)—meaning several voters might have made the same type of error, but there might be many different reasons why. While every best effort was made to present actual causal mechanisms, there is the possibility that some were not accurately identified. Future empirical studies would be required to confirm that the potential design deficiencies are in fact correct. Even with this uncertainty, it is still worth identifying in this chapter the major reasons why the failures might have occurred. Concrete design suggestions that have the potential to address the greatest usability impediments that keep voters from fulfilling their goals are required to improve future iterations of the system. While this method will not fix all of the system usability
problems, it can likely improve the vote casting and vote verification completion rates in the tested systems. This method will also help identify designs to avoid implementing in the next generation of e2e systems.

METHODS

The participants, materials, procedures, and general research design were the same as those described in the first part of the usability study (see Chapter 2). However in this second-part of the study, the main dependent variables were the behavioral errors and potential contributing design deficiencies linked to each vote casting and vote verifying failures. To collect the data, the experimenter directly observed each participant use the three e2e systems. If a participant could not cast a ballot or check on their vote, the experimenter took notes about how the participant deviated from the procedures that were required to complete each task. Then the experimenter noted system design deficiencies that might have contributed to the behavioral errors. System design deficiencies were identified by referring to resources that presented best practices of interface design (e.g., Johnson, 2010; Sanders and McCormick, 1993; and Tidwell, 2011).

RESULTS

For descriptions of the steps that users are supposed to go through to vote cast and vote verify, please refer to Chapter 1.
Helios

Vote Casting

As can be seen in Table 2, 41% of participants were not able to cast a vote using the Helios voting method. The most frequent reason that voters did not cast a ballot is that after encrypting their ballot, they did not login with their email account to both verify their eligibility to vote in an election and have their most recent cast ballot counted in the final election tally. Thirteen participants (35%) made this error. There are many design deficiencies that possibly contributed this particular error, each of which are discussed in the following paragraphs.

*Users who did not want to read and process all text displayed on the interface were not supported.* Figure 8 shows a screenshot of the user interface that the thirteen voters viewed when they announced to the experimenter that they finished casting their ballot. One potential design deficiency is that this interface does not support users who do not want to read large amounts of text. Users often engage in the practice of doing the least amount of work to try to accomplish their goal, even if it is wrong (Tidwell, 2011). In this study, it appears that participants did not cognitively process every piece of information shown on the webpage. Instead they likely saw the words, “We have received…,” and announced that they were done voting. At this point they thought their goal was fulfilled and did not put in further resources to make sure that it was actually achieved. The design of the interface did not support the user who wanted to quickly accomplish the task without putting in a great amount of effort to be more cautious. Some of the text was not worded clearly and hence was confusing. The most critical text was
not the most easily detected feature on the page. In addition information not related to that step was displayed prominently. For all of these reasons, the interface did not support actual voter behavior.

**Table 2. Errors and Likely Contributing Design Deficiencies that Led to Helios Vote Casting and Vote Verification Failures**

<table>
<thead>
<tr>
<th>Frequency</th>
<th>Error</th>
<th>Contributing Design Deficiencies</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Vote Casting</strong></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
| 13 / 35% | Encrypted ballot but did not log in with e-mail | • Users who did not want to read and process all text displayed on the interface were not supported  
• Log in information is not entered directly on page  
• Too many steps required to cast a vote |
| 1 / 3% | Did not press the cast button after a successful log-in | • Users who did not want to read and process all text displayed on the interface were not supported  
• Too many steps required to cast a vote |
| 1 / 3% | Refused to cast ballot due to concerns about errors | • Poor system usability degraded voter confidence |
| **Vote Verification** | | |
| 15 / 41% | Did not cast a ballot | • See above deficiencies |
| 6 / 16% | Performed wrong action | • No instructions provided for vote verification process |

**Notes:** This table does not include all observed critical incidents. It also does not include potential design deficiencies that contributed to partial verification failures, which are discussed below in the vote verification section.
Log in information is not entered directly on page. Another possible problem with the Helios system design is that the voters do not enter their email address and password directly on the page shown in Figure 8. Instead, voters are supposed to click the e-mail client icon, which takes them through another series of pages, and only then returns them back to a page that allows them to finally cast their ballot. Some participants could not figure out how to log in because it was not apparent to them that they should click on one of the icons to proceed. Of those participants who managed to figure out this step, two participants navigated away from the Helios voting system and could never find their
completed ballot again—ultimately giving up before casting their ballot. If participants could log in with their email credentials directly on the page shown in Figure 8, then there would be less of a chance that they would fail at logging in or become permanently lost in the process.

Too many steps required to cast a vote. Between the time the voter completes his or her ballot and then casts it, the person views at least five different screens (more if not already logged in). The developers of Helios were likely aiming to have the interface guide the user step by step in the prescribed order, which is generally considered to be a good interaction design. Perhaps, though, the steps with the accompanying information and actions required on behalf of the system user are not organized correctly. Particularly, the system does not strike a balance between the division of steps (and sub-steps) and the number of them, making the process so tedious that users look for the first instance that they might be done (i.e. seeing that the ballot was received), and then leave prematurely before the cast ballot button ever gets pressed. To prevent fleeing voters (Everett, 2007), it would be best to keep the distance short between the ballot completion page and the vote casting page in order to make the stepwise process the most efficient, and effective.

One participant failed to cast a ballot because he or she completed the login process, but on the very last page did not press the final vote-casting button. As mentioned above, this error likely occurred because there were too many discrete, cumbersome tasks required to vote—from ballot completion and review to multi-step encryption, logging in, and finally vote submission. Again, if some of the early steps were to be made invisible to the user (i.e., the systems still includes the security
mechanisms but do not require the voters to be knowledgeable of them), then the number of cumbersome tasks and the associated errors might be reduced.

Poor system usability degraded voter confidence. One participant refused to cast a ballot because she was concerned about errors. Specifically she was worried both that the system could be making mistakes and that she was not using the system correctly. For these reasons she wanted to talk to an election official to get help. Regarding her concerns about system errors, research should be conducted in the future to determine how to increase a voter’s confidence and trust in the system’s reliability and accuracy. As for her concerns about whether or not she was performing the correct actions required to cast a vote, addressing system usability issues would improve the participant’s confidence that she was using the system “correctly” and that her intended selections would be cast and counted accurately. Even though only one participant expressed these concerns, in a large-scale election the reported frequencies could be magnified so they should not be dismissed in this study.

Vote Verification

Ballot was not cast. In the vote verification process, 41% of participants were not able to check on their vote because they did not cast a ballot. In other words, after they thought they voted, they could not do anything else with the system because there was not a cast ballot to check on. To correct this problem, the system needs to be redesigned so every voter can both easily cast a ballot and recognize if they have or have not done so.
No instructions for vote verification process. Sixteen percent of the voters were not able to check on their ballots because they performed the wrong actions. This error likely occurred because voters were not given instructions on how to check on their vote, or even that Helios offered a means to verify that their ballot had been cast. Since participants were asked by the experimenter to verify their vote, they knew it was possible, somehow—an advantage study participants had that real voters would not.

There were many actions participants performed to achieve what they thought was a verification, when in fact it was not: participants viewed the review screen, googled their smart ballot tracker, and printed the smart ballot tracker. Some participants had no idea what to do. To reduce the frequency of these mistakes, the system should provide voters with simple, concise vote verification instructions both on the final vote cast confirmation screen and through the vote confirmation email. It would also be beneficial to provide inline instructions throughout the verification process.

Of the 43% of participants who were able to complete some form of vote verification, only half of these subjects completed verification procedures as outlined in the published papers describing Helios (e.g., Adida, 2008; Adida et al., 2009). In other words, only 22% of all the voters who participated in this study completed a full verification in which they checked their vote (associated with the smart ballot tracker that they recorded) on the online bulletin board. Voters who only partially verified did not do this—and instead relied on methods like viewing a vote confirmation email—likely because instructions on the vote verification process were not provided to the users. Even though this deficiency did not lead to a total failure, it is still noted because the
voters did not use the system as the designers intended in order to make sure that the system is working properly.

**Prêt à Voter**

**Vote Casting**

As can be seen in Table 3, when voting with the PaV voting system, 22% of participants did not scan in their ballot before placing it in the ballot box. It should be pointed out that the simple, single-use scanner was easily accessible. It automatically fed and scanned the ballot cards when they were placed into the feeder slot. In addition, the ballot box was placed off to the side of the vote casting station. There are several design deficiencies that might have contributed to this scanning error resulting in a failure to vote.

*No physical mechanisms on the ballot box prevented an unscanned ballot from being placed in it.* The ballot box was not designed to keep unscanned ballot cards (i.e., uncast ballots) from being placed in it. A design solution might be to place the scanner behind the ballot box opening. A voter would simply drop in their ballot for it to be both scanned and stored.

*System does not support prior voting mental model.* Related to the previous design deficiency, the system does not support participants’ general how-to-vote procedure for voting with a paper ballot. Specifically in past elections, voters might have completed a paper ballot and then placed it into a locked ballot box to cast it. Then in this study when participants used PAV to vote (which also requires filling out a paper ballot),
they likely referenced their previous model, which does not include the step of casting a ballot by scanning it into the system. Consequently they skipped this critical step, resulting in a failure to vote. To support these prior models of experience and knowledge, the PaV voting system should mirror the typical how-to-vote with a paper ballot procedure as closely as possible.

**Table 3. Errors and Likely Contributing Design Deficiencies that Led to PaV Vote Casting and Vote Verification Failures**

<table>
<thead>
<tr>
<th>Frequency</th>
<th>Error</th>
<th>Contributing Design Deficiencies</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Vote Casting</strong></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
| 8 / 22% | Did not scan-in ballot; placed it directly in locked ballot box | o No physical mechanisms on ballot box preventing an unscanned ballot from being placed in it  
  o System does not support prior voting mental model  
  o No inline instructions |
| 5 / 14% | Could not operate scanner | o Scanner not explicitly identified  
  o Operating instructions not available  
  o Design of scanner |
| 1 / 3% | Refused to detach candidates list and proceed with voting process due to being confused | o Numerous novel steps required to cast a ballot  
  o No inline instructions |
| **Vote Verification** | | |
| 14 / 38% | Did not cast a ballot | o See above deficiencies |
| 1 / 3% | Could not navigate to the verification website | o Requires use of the internet  
  o Website address too complex |

**Note:** This table does not include all observed critical incidents.

*No inline instructions.* PaV does not provide the voters with instructions when they are needed at key points in the voting process. Instead they are given long instructions on a single page before they start the voting process. This has the potential to
pose a large cognitive load on the user who is voting with a novel system requiring them to do unusual things in the context of voting, like tearing their ballot in half. A system improvement would be to give the voter clear, concise instructions at each step of the vote casting process. This design decision will also prevent voters from accidently shredding their instruction sheet and then having no idea how to proceed, which happened to one participant.

Five more participants did not cast their ballots because they could not identify and/or operate the scanner. Three potential design deficiencies have been identified.

Scanner not explicitly identified. Some voters had never used, or seen, a scanner before. This lack of previous exposure made it difficult to identify the single-function scanner and/or differentiate it from a single-function printer. See Figure 27. Labeling the equipment with both the word “scanner” and a visual icon might allow some people to at least identify the piece of equipment. Of those that could identify the scanner, some participants expressed that they were afraid of it because they have never scanned a document before. Even though a scanner is a relatively common piece of equipment, it should not be assumed that every voter knows how to use even a simple one. Thus there should be support in identifying it and operating it with confidence.
Operating instructions for scanner not available. Inline instructions for operating the scanner were not provided to participants. Especially for participants who never used a scanner before, it would have been helpful to tell them how to operate the scanner, even at a global level: e.g., “Insert pages in feeder slot for automatic scanning. No other action is required.”

Placement of scanner. Because of users’ strong mental models regarding the casting of a ballot into a ballot box, not a scanner, future designs might consider integrating these two steps into a single process. Particularly, if the scanner is placed inside the ballot box so that voters only have to drop in their ballot—a step that they already know how to do—and the scanner, hidden from view, completes the process, then the voters would not have to worry about “scanning.”

Numerous novel steps required to cast a ballot. One participant refused to detach their candidates list and proceed with the voting process because they were confused.
They did not understand why they would detach their selections from the people that they voted for and then shred them so there was not any record. Requiring the voter to perform so many atypical steps seemed to erode their confidence. It was also clear that they did not understand why they were doing what they were doing to make their votes private. To overcome this design deficiency in the future, the system should align as close as possible with typical how-to-vote procedures, and if a voter must perform a novel procedure, it should be easy to accomplish. It would also help the user if they understood why they were performing the action.

**Vote Verification**

Thirty-eight percent of participants did not verify that their vote was cast with PaV because they did not cast a ballot. Addressing system design deficiencies will reduce this rate of failure. One additional participant (3%) was not able to check on their vote because of two likely design deficiencies:

*Voter must use the internet and website address too complex.* PaV’s vote verification system required the participant to use the internet. However, the participant could not. In particular, they did not know in which field they should enter the website address; they kept entering it into the search engine’s search field, a common error among novice internet users (Sisson, 2014). Repeatedly, this participant also did not correctly type in the address, “votingresearch.rice.edu”—highlighting that the verification sites should have the simplest and shortest address possible.
**Scantegrity II**

**Vote Casting**

As can be seen in Table 4, 51% of participants were not able to cast their ballot with Scantegrity II. The same ballot box and scanner used in the PaV mock election were used with this system. For this reason, the possible design deficiencies are identical to those provided in the PaV section above and consequently should be referenced there.

**Vote Verification**

Fifty-one percent of participants were not able to verify their vote because they did not cast a ballot. Three percent, or one participant, could not use the internet to get to the website, the same error encountered by a PaV voter. Eight percent of participants were not able to verify that their ballot was cast because they did not record the ballot ID, which is required to login to the verification website. The following are the associated design deficiencies.

*Ballot ID not located in a prominent location on the ballot.* Per the City of Takoma Park, Maryland, Municipal Election, November 8, 2011 ballot, the ballot ID was placed on the front, lower right corner of this study’s ballot. A similar font and font size was also used. It is possible that placing the ballot ID in a location that was not prominent might have contributed to voters not being able to easily detect it, especially when there was a large quantity of information displayed on the ballot (i.e., election title heading, voting instructions, races with possible selections, revealed codes, and the unique ballot identifier).
Table 4. Errors and Likely Contributing Design Deficiencies that Led to Scantegrity II Vote Casting and Vote Verification Failures

<table>
<thead>
<tr>
<th>Frequency</th>
<th>Error</th>
<th>Contributing Design Deficiencies</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Vote Casting</strong></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
| 16 / 43%  | Did not scan-in ballot; placed it directly in locked ballot box | ○ No physical mechanisms on ballot box that prevented an unscanned ballot from being placed in it  
○ No inline instructions |
| 1 / 3%    | Could not operate scanner                       | ○ Voters did not have previous scanning experience                    
○ Operating instructions not available  
○ Design of scanner               |
| 2 / 5%    | No data available                               |                                                                     |
| **Vote Verification** |                                               |                                                                     |
| 19 / 51%  | Did not cast a ballot                           | ○ See above deficiencies                                             |
| 3 / 8%    | Did not record ballot ID                        | ○ No inline instructions  
○ Ballot ID not located in a prominent location on ballot  
○ Ballot ID is too long and complex, which deterred voter from making the effort to record it |
| 1 / 3%    | Could not navigate to website                   | ○ Voter did not know how to use the internet  
○ Website address too complex   |

Notes: This table does not include all observed critical incidents. It also does not include potential design deficiencies that contributed to partial verification failures, which are discussed below in the vote verification section.

**Ballot ID is too long and complex.** The ballot identification on every participant’s ballot was “HC-2016-11-08-420795502.” HC stood for Harris County, 2016-11-08 was the hypothetical date for the mock election, and 420795502 was the unique number associated with the particular ballot. This ballot ID was generated based on sample
ballots shown in published materials. Because the ID was so long and complex, it might have deterred participants from putting in the effort to record it. Worded differently, the perceived cost was too high for the return gain of being able to verify. Or perhaps participants did not recognize the benefit of being able to verify, so they did not want to transcribe the ID. On future Scantegrity II ballots, the ballot IDs should be as short as and simple as possible while still being secure.

Of the 38% of participants who were able to complete some form of vote verification, only 21% of these completed verification procedures as outlined in the published papers describing Scantegrity II (e.g., Carback et al., 2010; Chaum et al., 2008). Worded differently, only 8% of all the voters who participated in the mock election completed a full verification in which they correctly recorded their ballot ID, correctly wrote down all 27 codes, and looked at the vote verification webpage associated with their ballot. Since there is no data to confirm that participants actually compared all 27 transcribed codes to those displayed on the site, the full verification rate has the potential to be even lower than 8%. Voters likely only partially verified because too much of a mental workload and physical burden was placed on them. Specifically, after completing a ballot, participants had to transcribe by hand a long, complex ballot ID and small codes revealed inside each selected bubble. Then the subjects had to navigate to the vote verification website, and only if they wrote down all of the information correctly could they type in the case sensitive ballot ID and compare all the codes one by one. The other major problem with the Scantegrity II verification system was that there were too many opportunities for voters to make errors while transcribing and then comparing the verification information. Only 62% of all participants wrote down the
ballot ID correctly, and just 8% recorded all 27 codes accurately. The ballot ID and each code that had to be written down on the confirmation sheet is an opportunity for the voter to make a mistake. Then there are even more opportunities for errors when the voters type the ballot ID into the verification system and compare their recorded information with the information the system recorded. For all of these reasons discussed in this section, Scantegrity II had one of the weakest vote verification systems tested in this study. Even though the identified system design deficiencies did not lead to a total failure, it is still highly problematic because the voters did not use the system as the designers intended in order to check that their votes were recorded accurately.

Figure 28. An incorrectly completed Scantegrity II ballot and vote confirmation sheet
DISCUSSION AND CONCLUSION

This study’s findings indicate that the majority of the errors resulting in a total failure were caused by only a few contributing design deficiencies. This is positive because a small number fixes to the system design could solve the majority of the issues that caused voters to fail at vote casting. As for the verification systems, if the voters need to complete full vote verifications for security reasons, then the Helios and Scantegrity II methods should be reconsidered in future research. While every usability issue was not identified because the focus of this study was on fatal failures in which the voters could not achieve their basic goals of vote casting and vote checking, the data do provide direction for improving the systems to increase the likelihood that voters will fulfill these goals.

This study also reinforces that the errors were not a product of human fault. Rather the issues lie with the system design. Many human factors and engineering psychologists emphasize that systems can be designed differently to minimize the chance of events occurring in the first place—meaning errors can and should be designed out of the system (e.g., Norman, 1983; Shneiderman et al., 2009). While it is difficult to create error-resistant systems for diverse populations, it is especially important in voting. Every single person, despite age, background, or cognitive capability must be able to use the system to vote, otherwise election integrity will be threatened.

It might be argued that if these systems had been used in an actual election, some of the behavioral errors and consequent system failures could have been avoided if the voters had asked an election official for assistance. This may be true. Assistance, offered
at the right time, can prevent usability errors or rectify those that have occurred. While there are always officials present at polling sites, they should not be the first line of defense, and voters should not have to rely on them to successfully use the systems. There are high costs associated with voters asking someone to help them: voters might become frustrated and give up voting before asking for help, they might avoid voting in the future, their votes could be revealed to another person, officials might have to be trained to know how to address every possible usability issue, and there is the potential for voting times to increase, making lines at some polling stations even longer. For all of these reasons, systems should be designed so that any voter can walk up and effortlessly use them, without any assistance.

While the sample of 37 participants was diverse and adequately sized for behavioral research, watching this group vote did not identify every error that could lead to a failure. In order to develop an exhaustive list of low frequency errors, a different group of users from the same population will need to be used in future e2e research.

Another concern is that some of the usability issues identified are related to how the systems were setup for this particular research project. As mentioned in Chapter 2, if concrete system specifications were not provided, best practices of human factors were used. Even so, there were some behaviors, like voters failing to scan in ballots to cast them, which could not be anticipated since there were not existing data to know how voters would likely use the systems. Hence the value of this research is that these types of errors are now identified and consequently can be designed out of future e2e systems, which should be fully specified in detail.
Before concluding, it should be strongly cautioned that even if every possible system design deficiency identified in this chapter is addressed, the systems still might not be highly usable. For one, all usability issues were not identified and discussed. Only details about total failures were presented as opposed to every usability incident.

This research project also did not address if users who have physical disabilities would be able to use the e2e voting systems. It is hypothesized that Helios would be the most usable to this user group. Helios is implemented in an electronic format so existing software and hardware that allows the disabled to interact with computers and DREs can be paired with the voting system. For example, non-sighted voters could use screen readers, and voters with movement impairments could use equipment like jelly switches and puff and sip devices to help them vote. PaV and Scantegrity II would likely pose many unsolvable problems. How would non-sighted voters complete their paper ballots? Even if a device was available to read the PaV ballot to them, these voters would have difficulty knowing which pieces of paper to shred (i.e., the candidates lists) and which pieces to eventually scan (i.e., the completed ballot cards). Then the receipt issued to the non-sighted voters would be meaningless because it conveys spatial information displayed on a ballot card, which even a reading device would not be able to help them interpret. Similar challenges would be faced with Scantegrity II as non-sighted voters would not be able to read the revealed codes inside each selected bubble. Hence they would not be able to create their own receipt without assistance. Voters with other physical disabilities will likely have trouble tearing the PaV ballot in half, shredding the candidates list, and feeding the sheets into the scanner. In this study I observed a man who could not use the right half of his body resort to using his teeth to complete the
voting process. I also watched a woman who had problems standing have to sit down multiple times throughout the ballot completion stage of Scantegrity II. It took her a long time to make a selection, shuffle back and forth between the ballot and vote confirmation sheets, switch between writing devices (a decoder pen was used on the ballot and a pen was used to record the codes on the vote confirmation sheet), and write down the code that was revealed before going onto the next race. Since the task was so labor and time intensive for her, she kept having to sit down and rest. For reasons like these, research must take place to identify usability issues that the disabled will face and determine if and how the issues could be solved.

Another reason why all of the identified problems from this part of the study could be fixed and the systems still might not be usable is that the identified usability issues might have masked other problems that have yet to be identified. Once the problems are fixed, new problems might be found that could only now be recognized.

Usability problems might also persist if the proposed solutions do not adequately address the issues for a variety of reasons—from the wrong system design deficiency having been identified to the implementation of an insufficient solution. Accordingly, iterative development and testing should be used to move forward with fixing these systems so that they are highly usable.

In summary, the complex security mechanisms found in e2e voting systems do not comprise an insurmountable barrier to the development of usable designs. It is believed that by addressing the design deficiencies described here and implementing them using industry-standard methods like iterative design, the tested e2e systems can
become usable enough to be deployed in large-scale, actual elections—ultimately resulting in anonymous, secure, reliable, transparent, and accurate voting processes.
CHAPTER 4

Users’ Mental Models for End-to-end Voting Systems

INTRODUCTION

Voting security experts have designed end-to-end voter verifiable methods to improve the integrity of elections. In the process, they also wanted to enhance voters’ awareness of system security, make how the systems work more transparent, increase accuracy by having voters be active participants, and either increase voters’ trust in the systems that they use or eliminate the need for voters to trust that a system is working as it should. Not only have the system developers defined the operational improvements of each system, but they also presented a theoretical framework that sets their systems apart from non-e2e voting technologies.

By way of example, when the Helios system is described, it is touted as having the property of “unconditional integrity” (Adida, 2008, p. 336)—meaning the system is accurate and a tally cannot be faked. The developers of Helios thought this was the most important property that a voting system should have. They also believed that it was the one that was most important to other people, especially voters. Accordingly, the developers worked to ensure that the voters could trust the system itself since they thought it was better to trust a reliable, accurate system than potentially corrupt election officials.
While voters are supposed to trust Helios, Prêt à Voter (PaV) was designed in such a way that a person does not have to “trust” the system. Instead, the system “assures a high degree of transparency while preserving the secrecy of the ballot” (Ryan et al., 2009, p.662), meaning it provides evidence to the voter, throughout the voting and tallying processes, that proves the system is working as it should. Specifically, voters are issued a receipt that they can use to check online that their ballot was recorded as they cast it. Auditing teams can also check the decryption of votes without ever revealing specific voters’ choices. If voters or independent auditors notice that something is wrong, then fraud can be identified.

Scantegrity II “offers a level of integrity not found in conventional voting systems” (Chaum et al., 2008, p.45). To achieve this end, the system has features that allow the voters to check that it is in fact accurate. For example, when a voter makes a ballot selection, a unique code is revealed inside the selected bubble. This code, along with the ballot ID, can be recorded on a receipt. Voters can then choose to keep the receipt if they want to later verify that their votes have been cast as intended, all while keeping their actual selections private. Despite the inclusion of all these special features, the voter experience for Scantegrity II was intended to be identical to that of voting with an optical scan bubble ballot, except for the creation of the optional verification receipt.

But do voters who use the e2e systems associate concepts of increased accuracy, security, and transparency with the systems? All of these system elements require a voter to have at least *some* awareness of them and understand them, even if only at the most basic level. If voters do not, then the assurances and benefits of the systems are potentially lost. To make matters worse, voters might be frustrated with the deviations
from voting procedures that they are accustomed to or expect, or perhaps voters might not know what they are doing. All of these scenarios could result in negative consequences.

These issues are of concern not only to human factors researchers, but also to voting security experts. Ben Adida (2008) pointed out few people recognize that ballot casting assurance (when voters know that their votes were correctly captured) and universal verifiability (in which any observer can verify the correct tally of all votes) are two properties of the Helios voting system. Moreover, a system like Helios is not widely recognized to represent a major improvement over methods currently used in elections. Hence the only way to determine how voters think about these e2e voting systems is to gather empirical data. Only then can we know if the gap Adida describes truly exists.

The aim of this project was to understand the mental models users form after vote casting and vote verifying with the Helios, Prêt à Voter, and Scantegrity II methods. There was also a desire to determine if these models support the security concepts integral to the e2e systems. A glimpse into the mental models users have for these systems will allow several important questions to be answered: Do voters understand how the systems work? Do voters know why they are doing the things they do in order to cast a ballot or check on it after the election? And do voters find the systems to be more secure, trustworthy, or transparent than other voting methods like the paper ballot? Or do Voters think of e2e systems no differently?

To date, there has not been published research specifically on users’ mental models for any voting system. While there has been a very limited effort to study if users understand a specific system feature of an e2e voting system (e.g., why voters must tear
their ballot in half when voting with PaV, Winckler et al., 2009, see Chapter 1 for more information), the focus of the work was exclusively on assessing system usability, not on understanding user mental models.

In contrast, there is an extensive, diverse body of work on mental models in general. A single definition for mental models does not exist, as the term seems to be continually redefined across researchers and projects (Rouse & Morris, 1986; Tauber, 1988). For this reason several interpretations of mental models will be presented.

Rouse and Morris’s definition of mental models—based on Ramussen’s 1979 work—was “mechanisms whereby humans are able to generate descriptions of a system’s purpose and form, explanations of the system, functioning and observed system states, and predictions of future system states” (p. 351). Based on this definition, not only can a person describe what a system does and the form it takes, but the model also accounts for how the system actually works, what it is doing at a given time, and how it will act in the future. This means that the resulting mental models will account for both the frontend and backend portions of the systems.

Kieras and Bovair (1984) put forth a more straightforward definition for mental models as the understanding of how a device works in terms of its internal structures and processes. They also explained that mental models are not fixed, since they will change over time and across environmental contexts. There are multiple forms of mental models that might be expressed, dependent in part in how they are obtained—either through drawings, interviews, think out loud protocols, or a more formal diagramming method. Further, a mental model obtained by a researcher does not have to be perfectly
“accurate,” as this is an almost impossible feat; instead the utility of the model lies in terms of how it can be used to study a problem or system.

Donald Norman described mental models as evolving, incomplete models users have for systems with which they interact (1983). The models do not have to be completely accurate, but they should help the user in some way. He also observed that mental models are limited by a user’s background and experience using similar systems. User mental models are related to, but different from, the target system (i.e., the system a person is using), the conceptual model for the system as invented by the developers, and the scientist’s conceptualization of the users’ mental models. Worded differently, mental models are unique to the user of the system, but those presented in research are heavily influenced by how the researcher conducting the study acquires, analyzes, and presents their interpretation of the user’s mental model.

In this chapter, mental models for a voting system are defined using a combination of the previously described definitions: a mental model is a user’s understanding of what a system does, why the system does what it does, how one can interact with it, the expected responses to specific actions, system perceptions, and the ideas a user associates with the system.

This study aimed to capture participants’ mental models through drawings and interviews after using each of the e2e systems in a mock election. In turn, it was hoped that the mental models would reveal what users understood and did not understand about the systems. This information could then be used in future iterations of the e2e systems to decrease voter frustration, discouragement, or inability to act for any reason, while increasing their feelings of trust, security, and accuracy.
METHODS

Participants

Sixteen participants completed this study. These participants were eligible voters (i.e., 18 years old or older and residents of the United States) who lived in the Houston area. Subjects were recruited through two methods. First, eight participants were Rice University undergraduates recruited through the university’s subject pool. Each received credit towards a course requirement for their participation. Second, eight participants were recruited through an online advertisement. These subjects were paid $25 for their time. The mean age was 32 years, with a range from 18 to 65 years. Eight participants were male and eight were female. All had normal or corrected-to-normal vision. The mean number of national elections that these participants had previously voted in was 3.1, with a range of 0-12. On average, participants had voted in 4.3 other types of elections, (e.g., local, state, or school), with a range from 0 to 10.

Design

The study was a within-subjects design. Participants voted with each of the following systems: optical scan paper bubble ballot, Helios, Prêt à Voter, and Scantegrity II. Even though this study focused on mental models of e2e voting systems, participants were asked to also vote with a paper ballot so they would be able to directly compare the “enhanced” voter verifiable systems to the more traditional, basic method. For more information about these conditions and their related materials, refer to Chapter 2 that presents the usability study.
The main dependent variable was the mental model formed for each voting system after using it in the mock election. To capture the mental models, several methods were used. Participants drew on a piece of paper a representation for how they thought the voting system worked. They were also interviewed about their mental model drawing and asked further questions about security features, system accuracy, unnecessary procedures, things that they did not understand about the systems, their preferences, and how the e2e systems compared—especially with respect to security and auditability—to the paper ballot voting system. Basic demographic and background information was also collected.

**Procedures**

The study began by having participants complete an IRB approved informed consent form. The experimenter read to them study instructions and gave the participants a randomly assigned Republican or Democratic slate so they knew how to vote. This was the same directed voting procedure used in the usability study presented in Chapter 2. Subjects then went on to vote with a paper ballot in the mock election. After voting, the participants were asked to draw a representation or picture that described how the voting system worked—from the moment they were handed a ballot until the election outcome had been determined. They were then asked to verbally describe their mental model and answer follow-up questions that included a query about if the system had any security features, if they did anything unnecessary, or if they were unsure of what they were doing at any point while voting. Next they were asked to both vote and verify that they had voted with one of the three e2e voting methods (the order or presentation was randomly
assigned, and all orders were used). Even though vote verification is an optional step not required to vote with the e2e systems, participants were asked to check on their votes to see how they might go about doing this and if they understood the potential advantages. As with the paper ballot, they were then asked to draw their mental model and were interviewed about the system using the same types of questions. With the e2e systems, though, one additional question was included: “Do you think the system is more secure than a paper ballot voting method?” This question was asked to determine if the participants thought that the system was more secure than non-e2e voting methods. The participants repeated this portion of the procedure until they had voted with all the voting methods. Next they completed a final interview that covered topics like system preferences and how the systems differed with respect to accuracy and security. Last, the participants were debriefed, compensated, and thanked for their time.

RESULTS

A careful review of the participants’ drawn mental models led to the realization that almost every single participant represented on their paper only the steps required to vote, either generally, or with respect to the specific system. Out of the 64 ballots cast in this study’s mock election, 58 (91%) showed or listed out the steps a voter had to execute in order to have a ballot cast and counted in an election that would impact election outcomes. See Figures 29, 30, and 31 for examples. One participant drew a different symbol for each of the voting systems that expressed their impressions of the system. For the paper ballot method, the participant drew a snail symbolizing the slowness of the voting system. For Scantegrity II, the same participant drew a safe since he thought it was
secure. Another participant—for two of the four systems—drew a diagram of the respective ballot and the order in which it should be completed. But these two participants were anomalies. As a whole, when voters were asked to draw their mental model for the voting system that they just used, they expressed the how-to-steps required to cast a vote that would be later counted. As observed by Norman in his review of mental maps (1983), the drawings did not highlight every single step, the models were incomplete as they left parts of the system out all together (e.g., verification procedures), and they were inaccurate at times (e.g., some voters thought Helios included steps to make sure they were a human voter versus a computer system).

Figure 29. Two instances of participants’ drawn mental models for Helios
Figure 30. Two instances of participants’ drawn mental models for PaV
The mental models produced by hand did not vary a great deal across systems. It was not evident that participants had a deeper or more robust understanding of one system over others, or that participants accurately and fully understood any of the e2e systems (or even the paper voting method for that matter). What did change in the mental models associated with each system was the specific steps required to vote, equipment or
features unique to a system (e.g., completing a ballot on the computer when using Helios, the scanning in of ballots with PaV and Scantegrity II, the use of the special marking device that revealed invisible codes when voting with Scantegrity II, and the physical appearance of each system’s ballot).

**System Security and Accuracy**

**Paper Ballots**

Through an examination of both the drawn mental models and the interviews, it was revealed that some participants associated varying degrees of security features with the systems. With the paper voting system, only 13% of participants expressed any indication of security in the drawn mental model for the method. In contrast, during the interview 56% of participants expressed an opinion that there were secure features that they associated with paper ballot voting. The listed secure features from both data collection methods included a locked ballot box (38%) that would be tracked through a chain of custody (6%), manually counting ballots (13%) because they trusted their neighbors to do the task correctly, automated machine scanning and counting of the ballots (6%), and/or the fact that a change to the ballot would require physically altering it (6%). The presence of many of these security concepts, like the chain of custody, highlight that select participants do indeed have a more conceptually complex mental model for paper voting systems as they can explain in detail how the system works beyond rote procedures. However, overall few participants (6%) indicated this level of understanding. In contrast, forty-four percent of subjects thought there were no secure
aspects of a paper ballot voting system. Overall, it can be inferred from these reported percentages that the majority of participants either associate few security features with paper ballots or did not think they were secure in any sense.

**Helios**

For Helios, details about security were expressed equally in both the drawn mental models and the interviews. Seventy-five percent of the participants recognized that their ballots were encrypted by the system. Many of these subjects associated encryption with keeping their votes secure and/or private. At the same time, about 13% said they did not really understand what “encryption” meant, but knew it was a security feature. The smart ballot tracker (44%), voting on a computer (25%), the ability to audit the ballot (13%), and/or only being allowed to make one selection per race (6%) were also mentioned, but at a lower frequency than ballot encryption. Thirty-eight percent of the people seemed to have recognized that they had to confirm their eligibility to participate in an election after encrypting their ballot but before casting it. A minority of participants (13%) thought that this procedure was in place to make sure they were human and not a computer trying to cast votes to impact the election outcome. In actuality, this step was required to make sure a person is eligible to vote in an election, and allow them to cast and audit as many ballots as they want, with only the last one tallied in the final count. While every voter did not present a mental model saturated with features that are proposed to keep elections secure and accurate, it seems that most people did at least realize that there was something different about this system when compared to non-e2e voting methods.
Prêt à Voter

For PaV, the security aspect most often drawn and spoken about was the detachment from the ballot cards and subsequent shredding of the candidates list. Particularly 88% of the participants drew it or mentioned it. Some participants elaborated that this was to keep others from knowing how someone voted. Sixty-three percent of participants expressed that did not really understand why they needed to tear their ballot in half and shred a part of it. Nineteen percent indicated that they did not realize the candidates were randomly ordered for every participant and worried that their ballots could be reconstructed to determine how they voted. Twenty-five percent felt that scanning the ballot cards would help keep an accurate record and count. Seventy-five percent indicated the system issued a printed receipt that could be used to go online to make sure the system counted their vote. Thirty-one percent explained that the verification system showed their actual selections on the ballot card. The take-away for PaV is that some participants recognized the features that kept the system transparent, secure, and accurate, but for the rest of the participants it left them confused and without understanding that the system had potential benefits to voters. Out of the three voting systems, PaV’s actual security mechanisms seemed to be most recognized.

Scantegrity II

Security and accuracy features voters associated with Scantegrity II included the vote selection confirmation sheet (19%) in which the voters recorded their unique ballot ID (25%) and codes (69%), the special marking device used to complete the ballot (44%), scanning of the ballot (44%), the storage of the ballots in a locked ballot box (13%),
and/or the booth participants stood in while completing their ballot (25%). There were points of confusion surrounding Scantegrity II’s perceived security features. Thirteen percent did not know why they had to use a special marking device. Thirteen percent complained that they could not check their actual votes online, as they were only shown codes. And 6% said that the ballot did not keep their selections private because anyone could look at it and see which bubbles they filled. What is interesting about this list is that some of the items are those that were purposely included by the designers, such as the marking device and confirmation sheet. Others, however, are physical attributes of the equipment that were not central in the developers’ conceptual model, like the locked ballot box and voting booth. Even if these items were inferred or briefly mentioned in published papers, they were not integral to what sets Scantegrity II apart from other systems (and it should be noted the same type of ballot box and voting booth was used with every system in this project). Even though the occasional voter mentioned the voting booth or locked ballot box in their PaV mental model, these items did not necessarily dominate their list of security features associated with the system.

**System Security Comparisons**

Participants were asked if they thought each e2e system was more secure than the paper voting method. Sixty-nine percent of participants said Helios was more secure. They thought the system was more secure than paper because voting took place on a computer, the ballot was encrypted, and no humans were involved who could make mistakes. As for the 31% of participants who said Helios was not more secure, they were concerned about hackers, computer glitches, and the fact that they had no idea what the
system did with their data (i.e., there was the potential for the computer to do one thing but tell them another).

Sixty-three percent of participants thought PaV was more secure than the paper method. Evidence to support their opinion included automated record keeping and tallying, a candidates list separate from the ballot selections to keep votes anonymous, and online verification. The 38% of participants who said, “no” cited that the candidates list was in the same order for everyone so votes could be reconstructed, the removal of their candidates list was keeping them from seeing who they voted for when viewing their receipt at home, the system was too complicated, and there was a possibility of hacking since the votes were stored and tallied on a computer.

Half of the participants thought Scantegrity II was more secure than voting with a paper bubble ballot. When using Scantegrity II to vote, a group of participants said there was less chance for human error due to automated ballot scanning and vote tallying. The special marking device and online verification system made it more secure. And one participant did not know what the codes were for, but they made him feel “psychologically secure.” The other half of participants did not think Scantegrity II was more secure because of things like not knowing the benefits of the codes or which codes corresponded to which selections.

When participants were asked which system used in the study was the most secure, half of them identified Helios (see Figure 32). Participants justified their selection by saying that it was automated, encrypted, online, and they just “knew” it was more secure. With that being said, there was no statistical evidence to support that subjects felt there were differences in security across the voting systems, \( \chi^2 (3, N = 16) = 5.5, p = .14. \)
When participants were asked which system was the least secure, 47% said the paper voting method was least secure and accurate (see Figure 33). With this system, the problems included people handled the papers and counted the votes, there were many opportunities for fraud to be introduced, and anyone could see how someone voted. It should be noted that there was no statistical evidence to support that subjects distinguished one system to be less secure than the others, $\chi^2 (3, N = 16) = 5.75, p = .12$. A posthoc power analysis revealed the sample size was not large enough to detect effects if they were present, hence future research should address this issue.

![Figure 32](image_url)

**Figure 32.** Frequency as a function of voting system type for the system participants indicated was most secure

To summarize the security findings, voters recognized security and accuracy features of the system, and these features varied across the four systems. But participants
neither thought e2e voting systems were more secure or accurate than the traditional paper ballot—implying there is a gap between the actual integrity of e2e systems and the perceptual integrity of the systems.

**Figure 33.** Frequency as a function of voting system type for the system participants indicated was least secure

**Unnecessary Steps and Procedures**

Participants were asked if they thought they performed any unnecessary steps or procedures while voting with the systems. This question was asked to determine if there was a procedure or mechanism central to the voting system that they did not understand, or that they felt was superfluous to the core act of voting. The logic was that if they did understand why a novel procedure or feature was implemented, then they would not think it was unnecessary. But if they thought it had no purpose, then it could be inferred that
they did not fully or accurately understand how the system worked or the benefits of it. For Helios, 38% of participants said there were unnecessary steps, system features, or equipment. Their explanations included being an active participant in the ballot encryption process, being issued a smart ballot tracker, logging into their e-mail account before casting their ballot, and the high number of steps. Seventy-five percent of participants said PaV included unnecessary procedures and components: detaching and shredding the candidates lists from the ballot cards, shredding the candidates lists before being issued a receipt, and the option to verify online. Fifty-six percent of participants felt Scantegrity II had unnecessary elements, which included the codes revealed by the marker, online verification, having to write down so many codes, and the number of steps required to cast a ballot. About half of participants felt there were unnecessary system components and vote casting procedures. Importantly, many of the cited components were the very features—according to the system developers—that make the systems accurate, transparent, secure, anonymous, and audible. This indicates that the system developers have a significantly different conceptual model for the e2e systems than the actual voters. It could be argued that these different perceptions are actually advantageous for voters, because they could increase the pleasantness of voting with the systems or promote future system use. This phenomenon could rightly be called the “ignorance is bliss” hypothesis—some voters do not know, and do not feel that they need to know, every nitty-gritty detail of the system as long as it always works.
Procedural Uncertainty

Participants were asked if they were unsure of what they were doing, or why they were doing it, at any point while using the system to vote or verify their vote. The reason for asking this question was to identify both the aspects of the system that the participants did not understand and the how-to-vote steps that were not clear. Forty-four percent of participants said they were uncertain as to what they were doing while using Helios. Some participants did not know what encryption was and/or were unsure how to encrypt their ballot. Others did not know what to do with their smart ballot tracker, how to print the webpage that showed their smart ballot tracker, or how to cast their completed ballot. Fifty-six percent of subjects said that they were uncertain at some point while using PaV. They did not know if they really should tear their ballot in half and shred a part of it, why they needed to remove this candidates list, or if the page order mattered when they scanned in their ballot cards. The highest uncertainty response rate was for Scantegrity II, with 75% of participants having said they were unsure. Participants said there were too many directions for the system. They also noted that there were many things they did not know what to do with: e.g., codes, the special marker, or how to verify. Some had trouble operating the single-use scanner, which only required the participant to insert the ballot into it as it then automatically fed the paper through. The majority of these confusion points relate to executing how-to-vote procedures, with the minority dealing with the reason why a step was required to vote or check on their vote. If the participants’ mental models were adjusted in these areas so that they understood how the systems operated and the resulting system states, perhaps confidence and performance could be boosted. At
the same time, readers should be reminded that comprehensive, detailed conceptual models should not be required to vote.

**System Preferences**

When asked which of the four methods participants would prefer to use in a real election and why, 38% preferred Helios (see Figure 34). The participants thought Helios was straightforward, fast, secure, familiar since it required the use of computers, and advantageous that it only required voters to figure out how to use one piece of equipment. Yet, there was not evidence to support a difference in proportions of responses across systems, \( \chi^2 (3, N = 16) = 3.5, p = .32 \). When asked which of the four methods participants would least prefer to use in a real election, participants least preferred Scantegrity II, with 38% (see Figure 35). Participants said the system was “annoying,” it was hard to read the codes in the bubbles, they could not fix mistakes or change a vote on their ballot once it was marked, the confirmation codes were an annoyance to record and write down correctly, some people knew there should be codes revealed but they could not find them (since they used a pen vs. the special marking device) and hence could not verify their votes online, and they overall felt the system was confusing. Again, there was no evidence to support a difference in proportions of responses across systems, \( \chi^2 (3, N = 16) = 3.5, p = .32 \). Like earlier in this section, the posthoc power analysis revealed the sample size was not large enough to detect effects if they were present, so future research should address this issue.
**Figure 34.** Frequency as a function of voting system type for the system participants would most prefer to use in a real election

**Figure 35.** Frequency as a function of voting system type for the system participants would least prefer to use in a real election
In many ways, these portions of the participants’ mental models, made-up in part by their experiences with them, highlight the best and worse features of all of the systems. Participants want a system that is easy to use, efficient, accurate, and secure. In contrast, subjects wanted to avoid confusing, lengthy procedures, and not knowing how to vote and verify, even though some relayed that they could do this. What is more, the most preferred system is a great improvement in terms of election integrity over the least preferred system. This alignment is beneficial because it shows that voters recognize strengths in e2e systems like Helios, even if for not for the same reasons the system developers find them to be better.

**DISCUSSION AND CONCLUSION**

The drawn mental models of end-to-end voting systems generated in this study outlined each system’s how-to-vote procedures. They did not show any of the inner workings for the systems or explain how they operate. But when interviewed, it became clear that a subset of the participants could describe the presence of system features that made the systems secure, accurate, and transparent—setting them apart from the paper ballot. At the same time, many participants did not view these newer e2e systems as being more secure or accurate than a paper-based voting method. In the mental models, no participant could describe with complete accuracy or in any detail how the systems function or why they do what they do. Thus participants have mental models for the e2e voting systems, but they are incomplete, as they do not account for the entire system (e.g., they do not understand the system cryptography, and some voters do not recognize that the systems were designed to protect voter privacy). Participants’ mental models do
not significantly change depending on the voting method used and at times are inaccurate. Also these mental models are different from how voting security experts conceptualize these systems and seem to expect others to think about them.

The nature of the mental models found in this study are likely a product of the fact that many participants’ only exposure to these e2e systems was a result of participating in this study. The subjects read lengthy instructions told them how to vote with a method, then they voted with the system by completing a series of steps using the provided equipment. They were never told what was special about the systems, nor did they likely have any conceptual framework for the e2e system. Therefore, it is unsurprising that the participants could not do anything else but describe what they did to vote and the novel features and procedures that they encountered. Without access to the researchers’ concepts behind the systems there was no way for the participants to articulate these concepts.

Importantly, this is not a limitation to this study. Voters do not generally get to choose what kind of system they use to vote (except that they may be able to choose to vote by mail or use the one at their designated polling station). Voters cast a ballot with the system that is used at their assigned polling location or vote center; these systems are chosen by experts and placed into polling stations at the convenience of the election authorities. As such, voters in real elections would encounter the same predicament as the participants in this test and would likely end up with similar mental models for the system they just used. This disconnect must always be considered when forming assumptions about what the voters know about the systems and how they view them.
Future research is required to determine how the participants’ mental models can be enhanced or altered to avoid moments of confusion, uncertainty, and negative affect, and in turn make the systems trustworthy and truly transparent. One area for further exploration might be to examine whether the implementation of cues to indicate the presence of a specific security feature and an inline explanation of why that feature exists might help voters form a better, more complete model without causing them to slow down. For example, if Helios tells voters not only that their ballot is encrypted, but in a sentence or two why encryption is important, then perhaps users might not feel uncertain about the benefit of encrypting their ballot. Or if Scantegrity II indicated to voters that each code is unique, associated with the selections on their particular ballot and no other, and are used to keep their votes completely private yet trackable, then perhaps voters will recognize their value.

Training and teaching voters about the system before voting with it is not an ideal option. Aside from the obvious logistical difficulties, voters, particularly technologically phobic voters, might simply refuse to vote if they believe that a complicated system is required. Voting must be effortless. Voters might also have the potential to misunderstand what is taught to them about the system. If this happens, the information could interfere with their performance and result in negative outcomes. Alternately, voters could learn about a particular e2e system, show up at their polling place, and find they need to vote with a totally different system. Consequently, figuring out how exactly to communicate information that could enhance and clarify users’ mental models is integral to the user success of these systems. That said, this information must be delivered to them in an artful way because this study’s participants already complained that the
instructions were long and confusing. For this reason, simply expanding the instructions that are presented at the beginning of the voting process is not a good strategy.

While it would be wonderful to be able to make certain that voters understand a system that they use to vote in order to increase their confidence in the voting method and recognize it as an improvement over other methods, voters should not be required to have a comprehensive, detailed mental model for a system in order to use it successfully. Think about the telephone; almost every person can pick up a landline and dial a phone number. Few people can probably describe the operational nature of the phone system (Kieras & Bovair, 1984). Similarly, voters just need to be able to get through the how-to-vote sequence of votes that allows them to easily cast a ballot as they intend. They should not have to understand the system, or think it is accurate and secure, to be able to cast a vote that is counted.

One limitation of this research study on voters’ mental models of e2e systems is that there are many sources of potential bias introduced into the models. The mental models presented in this paper are the author’s conceptualization of these participants’ mental models (see Norman, 1983). Consequently, the voters’ interpretations are directly impacted by the training, background, and approach that was used to study the problem. The models were also impacted by the methods used to collect the models because participants might not be able to fully express their mental model through any singular method. These are limitations that every mental model study faces, and there is still utility in studying them (Kieras & Bovair, 1984).

Another limitation with the study is the sample size and how the participants were recruited. The sample size is small, and is not fully representative of the population since
half of the subjects were Rice University students. Nevertheless, even the students are real voters, and offer a glimpse into how people might think about the system.

There are many avenues for future research. For one, researchers could try to capture mental models through other methods such as think-out-loud protocols or alternate interview formats. Doing so would surely reveal aspects of the voters’ mental models not revealed in this study. Second, methods of altering mental models and how subsequent voting performance changed should be studied. Finally, the same techniques as those used in this chapter should be employed after voters use future iterations of these same systems to understand how the models change as the systems are manipulated.

In conclusion, voters do not have comprehensive, conceptual mental models for any of the tested voting systems, including paper ballots. However, they do have mental models that highlight the steps required to vote in order for their voice to be heard in an election. While this study is not the final word on the matter, it does provide a glimpse into how voters are actually thinking about the systems as opposed to how others expect them to think about them.
CHAPTER 5

To Verify or Not to Verify, That is The Question: Voters’
Anticipated Use and Expectations of Voting Verification
Methods

INTRODUCTION

How confident can someone be that a vote he or she cast in a presidential election was actually counted in the final total that predicted the winner of the race? Any individual should not be too confident. A lot may have gone wrong. Election officials could have unintentionally miscounted votes. Ballots might have been intentionally lost. Votes for specific candidates could have been added into the total. The system used to vote could have been designed so that voters were only able to select one candidate. Electronic machines might have reported fabricated numbers for the election total. In other words, there are many opportunities for fraud and malicious attacks on voting systems (Miller, 2007; Lindeman & Stark, 2012; Ryan & Peacock, 2005) because most of the currently used technologies are not designed to ensure that these kinds of acts do not occur. This is extremely problematic because if a voter’s cast ballot is not accurately cast and counted, the democratic process is in jeopardy. In response to these threats, computer scientists developed auditable voting methods that allow voters to check on their cast votes, which have been described throughout this dissertation.
As described in Chapter 1, every e2e system has been designed to allow voters to check on their ballot after casting it. With that being said, an important piece of the equation has been almost entirely ignored: Do voters actually want to verify their votes? If voters are already highly confident that their votes are accurately counted, then maybe they will not be interested in going through the additional steps to verify. Or maybe they are confident in their voting method, but they just do not care about checking on their vote, either because it will take too long or simply because they are overconfident. In all of these instances, a verification system might not be meaningful to and/or desired by the voter.

There are also other critical questions that have been left unanswered regarding the design of the verification systems. For instance, if voters really do want to verify their votes, does the form of verification match the options the voting researchers are providing? See Figure 36 for a screenshot of each tested system’s verification page that it displays to voters. Perhaps voters want to see something else on the verification webpages, meaning the systems should be adjusted? These are important questions to ask, because a verification system might not be used if voters do not think it provides them useful information or the specific type of information that they want to see.
Figure 36. Screenshots of Helios’ (shown left), PAV’s (shown center), and Scantegrity II’s (shown right) vote verification pages that may be viewed by voters if they choose to check on their cast ballots. The PAV and Scantegrity II pages are specific to this study.

There are many reasons why a person might not want to check on their ballot after voting—from the task being perceived as too cumbersome to general disinterest. It would be problematic, however, if enough voters do not complete the verification process carefully because the additional security and accuracy afforded by the verification step only works if sufficient numbers of voters actually perform the post-voting verification (Benaloh, 2007). This point adds further impetus for studying the issue.

The research described in this chapter is designed to begin answering questions regarding vote verification. Three fundamental questions that have yet to be answered by the voting research community were explored: First, how confident are voters that their votes will be cast and counted correctly using current voting methods? This question is asked because if voters have very high confidence levels then they might not verify that
their ballot was cast. In contrast, if their levels are lower, not only could future
verification behavior be impacted, but it would bolster the need for future research to
determine if confidence can be improved with vote verification system use.

Second, are voters interested in being able to verify their votes after leaving the
polling station, and would they actually perform this verification? Even though projected
behavior can be different than actual behavior (Berendt et al., 2005), it would be worth
knowing if voters have a desire, or not, to check on their votes. If voters already indicate
that they will not verify, then maybe whether or not verification systems should be used
would have to be considered. It might be that the verification systems would have to be
designed strategically to motivate voters to use them so that there are enough voters to
detect fraud and inaccuracies given the particular election’s margin.

Third, what forms of verification would voters want or expect to see on the
verification website? This question was asked to determine if current e2e systems match
voters’ expectations. If expectations and realized forms do not match, then the systems
should be modified.

By answering these preliminary questions, the void in the voting research can be
filled. As a result better design decisions regarding the implementation of voter
verification mechanisms can be made so that voters will be more inclined to use them.
METHODS

Participants

The participants were 767 registered voters in Harris County, Texas. To recruit respondents, pollsters called a pre-compiled list of registered voters in the county and asked them if they would like to participate in a pre-election poll. Respondents were not compensated for their time. The ages of those who participated in the study ranged from 18 (the minimum age to vote in the United States) to 104 years, with the mean age being 58 years. 356 participants were male and 411 were female.

Design

The study was a pre-election poll, composed of 35 questions. The poll was conducted by telephone over a two-week period, starting about three weeks before the 2012 Presidential Election.

35 questions made up the pre-election poll. The questions addressed various voting and election topics, of which three directly related to this research project. Please see Table 5 for these questions.

It should be noted that the responses for the preferred formats of verification question were developed so that all possible forms of verification would be represented, some of which are already implemented in end-to-end voting systems. “A digital image of your completed ballot” refers to PAV’s scanned ballot card images. “Confirmation that your ballot was cast and counted in the total” is essentially Helios’ form of verification as a voter’s smart ballot tracker is displayed on an election’s online bulletin board. “A
confidential PIN number that is associated with each candidate you voted for” aligns with Scantegrity II’s system that displays the code inside each selected bubble on a ballot. “A list of names of all the people that you voted for” was also included as an option—even though it is not feasible in U.S. democratic elections in which votes must be anonymous; in previous voting studies run at Rice University, participants mentioned that if they verified, they would want to see a list of names for which they voted.

Table 5. Research questions (and their response sets) that are of central concern to this project

<table>
<thead>
<tr>
<th>Question</th>
<th>Responses</th>
</tr>
</thead>
<tbody>
<tr>
<td>How confident are you that the vote you will cast/have cast in this year’s Presidential Election will be counted accurately?</td>
<td>Very confident, confident, somewhat confident, not confident, or unsure</td>
</tr>
<tr>
<td>If you could go home after you cast your ballot in this year’s presidential election and confidentially verify on the internet that your vote was cast and counted as you intended, how likely is it that you would do this?</td>
<td>Very likely, somewhat likely, unlikely, very unlikely, don’t have access to the internet, or don’t know</td>
</tr>
<tr>
<td>Which of the following would you want to see on a website that allows you to verify your vote?</td>
<td>A digital image of your completed ballot; a list of names of all the people that you voted for; confirmation that your ballot was cast and counted in the total; a confidential PIN number that is associated with each candidate you voted for, and you are the only person who knows what each stands for; none; or don’t know</td>
</tr>
</tbody>
</table>
Procedure

Registered voters in Harris County were called on the telephone. They were asked if they would like to participate in the IRB approved survey about November 2012’s Presidential Election. If they gave their approval to proceed, they were then asked up to 35 questions. Fewer questions were asked if material was not applicable to the respondent. For example, if a voter never votes straight ticket, then they would not have been asked questions like “For which party did you vote/will you vote a straight ticket?” After all applicable questions were asked, participants were thanked for participating in the survey.

RESULTS

Confidence in System Accuracy

The first question was, “How confident are you that the vote you will cast or have cast in this year’s Presidential Election will be counted accurately?” As can be seen in Figure 37, more than half of the respondents were confident to some degree that their vote would be accurately counted in the election. 336 people were very confident and 186 were confident. In contrast, about 28% of those polled were more skeptical: 154 were somewhat confident, 60 were not confident, and 31 were unsure. It should be noted that the respondents might have previously used or will use the Hart Interactive eSlate machine. That system is the voting system used in Harris County where the participants are registered to vote, and user interactions with this system might have impacted these findings.
Figure 37. Number of responses as a function of voter confidence level in cast votes being counted in the 2012 Presidential Election totals

Desire to Verify

The second question was, “If you could go home after you cast your ballot in this year’s Presidential Election and confidentially verify on the internet that your vote was cast and counted as you intended, how likely is that that you would do this?” Figure 38 shows that 496 respondents (65%) thought they would likely verify their vote through an online vote verification system, while 211 (27%) were unlikely to do so and 24 (3%) were not sure what they would do.
Figure 38. Number of responses as a function of likelihood that voters
would check-on their votes through an online verification system

Verification Format

The last question was, “Which of the following would you want to see on a
website that allows you to verify your vote?” The response set included, “a digital image
of your completed ballot,” “a list of names of all of the people that you voted for,”
“confirmation that your ballot was cast and counted in the total,” “a confidential PIN
number that is associated with each candidate you voted for, and you are the only person
who knows what each stands for,” “none,” and “do not know.” Respondents were
directed to choose as many options as they would like. Out of all the participants, 195
(23%) did not want any form of verification. 199 people (23%) wanted a general
confirmation that their vote was cast and counted in the election total. 152 respondents
(18%) wanted to see a unique, confidential PIN representing their vote. 117 people (14%)
wanted to see a list of all of the candidates and/or propositions for which they voted. 94 respondents (11%), wanted to see a digital image of their cast ballot. 77 people (9%) did not know what they wanted to see, and 18 did not respond to the question. See Figure 39 for a summary of these results.

![Figure 39. Number of responses as a function of preferred form of vote verification.](image)

If the responses to the third question are examined by only including respondents who would be likely to verify their votes after voting, essentially the same preferences were expressed with the most popular option still having been the general confirmation that the ballot was cast and counted (179 people), followed by the PIN option (140) and
list of selections (108), and the least popular having been the digital image of the cast ballot (90).

**DISCUSSION AND CONCLUSION**

The data revealed that only 68% of Harris County registered voters were confident that their votes would be cast and counted correctly, while 29% were somewhat or not confident. It is troubling that voters were not more certain that the voting method they use to vote would work like it should. This drop in confidence may simply reflect the steady decline of trust in the government that has been seen since the 1960s (Hetherington & Husser, 2012; PEW, 2010), or it could be a direct reflection of distrust associated with electronic voting systems (Oostveen & Van den Besselaar, 2004).

The wide variation of current voters’ confidence levels adds impetus to the direction the voting security researchers are moving with voting system development, which provides a means for voters to verify their votes after casting a ballot. By giving voters a way to track their vote to the end, it requires them to have less blind faith in a system because they know—through evidence—it is doing what it should be doing and reaffirms that the system is secure from attacks and manipulations. At the same time, it seems unreasonable to ever expect all voters to be confident that their vote will be counted accurately. Due to individual differences in voters (Rothschild & Houston, 1980), there is always the potential for voters to be skeptical or untrusting no matter how secure or accurate a system is proven to them to be.

In future research, it would be important to see if voters are more or less confident after using the recently developed end-to-end voting systems. Since voters have a way to
check that their votes have been counted as cast, it is plausible to assume that this would enhance voters’ confidence in the election. On the other hand, there is the potential that these kinds of verification systems might actually decrease their confidence in the system. The added verification procedure might remind them of all of the malicious things that can happen to their vote and make them question the trustworthiness of the verification process itself, ultimately reducing their overall confidence in voting. Either way, the ultimate goal should be to increase voter confidence, and end-to-end verification mechanisms might just be the solution to do that.

The data from this study highlight that the majority of registered voters (65%) wanted the ability to check on the internet that their vote has been cast and correctly counted. This finding confirms that voters are potentially interested in actually verifying their vote—meaning it is not something that only voting system developers think that people will want to do. In other words, the security researchers’ model of voter behavior is aligned with voter’s preferences in this particular design decision.

Voters wanting to verify their vote and actually performing the procedure to check on their vote are certainly two different issues. Verification systems do not rely on every voter checking the correctness of their counted ballot. Even if some voters are not interested in undertaking the verification process, those who do actually carry out the verification procedure will help improve the integrity of elections (Benaloh, 2007). However, with 65% having indicated some form of interest, this would seem to suggest that there will be sufficient numbers of voters who would actually perform this check in a real world implementation. At the same time, recent research found that only 2.8% of voters who cast a ballot with an e2e system went on to verify their vote online (Delis et
al., 2014). If actual usage rates turn out to be this low across systems, then it would have to be determined if enough voters are participating in the vote verification process to preserve the integrity of elections. An alternate explanation might be that voters want the ability to audit the voting system, but they might not actually do it unless there is a concern.

Since the majority of voters wanted some form of verification, the final, and perhaps most important, piece of the puzzle is determining what form that verification should take. About a quarter of respondents said they did not want to see any form of verification. This correlates almost perfectly with the percentage of voters who said that they were uninterested in verifying their vote (27%).

As for those that specified forms of verification, about a quarter wanted a general confirmation that they voted without giving any specifics. It shows that a simple implementation where an email is sent to every voter after voting to explain that their vote was cast, or a general ballot tracking number displayed on a community webpage, might be sufficient to fulfill some people’s expectations (both of these forms are currently used in the Helios system; see Adida, 2008). This verification form contrasts with the more detailed verifications provided by end-to-end systems like Scantegrity II and PaV (Chaum et al., 2008; Ryan et al., 2009). Of the respondents who wished for a more elaborate verification format, 11% wanted to see a partial digital image of their ballot, (as implemented in Pret-a-voter) and 18% a unique PIN association with each ballot selection. (the crux of Scantegrity II’s verification system). Both of these verification methods have been implemented in one or more of the verifiable voting
systems described in the literature, and the verification systems are not at odds with the preservation of voter anonymity or selection privacy.

In contrast to the other vote verification formats that preserve voter anonymity, 14% of the respondents wanted to see a list of everyone whom they voted for. Some of the respondents who desired digital images of their ballots might have also interpreted that question to include the names of the candidates for whom they voted, and not just a representation of the ballot selections. Accordingly this percentage is likely higher. These forms of verification would be problematic, because providing such a list or image reveals whom voters voted for—breaking privacy and allowing such practices as voter coercion and vote buying to more easily occur. While it is understandable that voters want to see a complete record of their vote, it simply cannot be done in the United States where anonymity is considered a cornerstone of the voting process. Further research is needed to see if there might be verification formats that would satisfy these users while maintaining the anonymous nature of the vote.

So what can be made of this study’s findings? When it comes to verification expectations, one size does not fit all. Even if the most popular option was provided—a general confirmation that a ballot was cast and counted in the total, there would be a substantial number of voters who would not be satisfied that their vote was counted as cast. At the same time, if unique PINs or codes associated with each selection are provided to voters, then another subset of voters will not be happy because their verification scheme provides too much information—thus complicating the voter’s ability to determine a successful vote. Perhaps voters should be given the option of a general confirmation (like a simple email after voting), along with a more information rich form,
where partial images or unique PIN codes are provided. In this case, voters could decide on a case-by-case basis the verification option they would like to choose, and if they want to verify at all. Of course, this kind of implementation would add further complexity to the already complex secure voting systems. And to make the matter even more complicated when designing the verification systems, there is not a single preferred option for verification that stands out—regardless if users have interest in verifying their votes or not. For this reason, more research is needed to determine how to maintain security and anonymity while still fulfilling the expectations of all voters.

This research should serve as a reference point for future system development and research. In many ways, the research has raised more questions than it has answered, and significant research into voting verification systems needs to be conducted in order to construct methods that adequately address the needs of security, anonymity, and usability. Voting system developers will need to consider all of this research when designing the next generation of secure voting systems, so that voters will be confident that their elections are secure and their intended votes have been counted accurately.
CHAPTER 6

General Discussion & Conclusion

Overview of Research

Through this research project, baseline usability data was collected for three end-to-end voting systems—Helios, Prêt à Voter (PaV), and Scantegrity II—which are representative of verifiable voting systems. An attempt was also made to identify why behavioral errors that led to vote casting and vote verification failures occurred, understand how users thought about the e2e methods, and identify their expectations for forms of vote verification. Three studies were undertaken to this end:

*Study 1, Part 1: Usability of Voter Verifiable End-to-end Voting Systems: Baseline Data for Helios, Prêt à Voter, and Scantegrity II.* The data indicated that the tested e2e methods were difficult to use. Across systems, only 58% of all ballots completed in this study were successfully cast. It took voters longer to cast a ballot with these systems than with previously tested voting technologies. Satisfaction levels for the three e2e systems were comparable or lower than those used in the previous studies. Even fewer participants could verify their votes. It also took them about the same amount of time to verify their vote as participants took to vote using more traditional voting methods.

*Study 1, Part 2: From Error to Error: Why Voters Could Not Cast a Ballot and Verify Their Vote.* The reasons why voters failed to cast a ballot and later verify that their
ballot was cast were identified in this part of the study. The most frequent Helios vote casting error was a failure to log in with an email address, a step after ballot encryption but before ballot casting that is required to proceed with the voting process. When voting with PaV and Scantegrity II, participants most often failed to complete the task because they did not scan-in their ballots. In the vote verification process, most participants were not able to check on their votes—through any means—because they did not cast a ballot in the first place. Likely contributing design deficiencies for all of the systems were also assessed. A careful examination of these revealed that a few design decisions were driving the majority of the behavioral errors that caused voters to fail at completing their goals.

*Study 2: Users’ Mental Models for End-to-end Voting Systems.* Voters’ mental models of the tested e2e systems were dominated by how-to-vote procedures. Some participants were aware of novel steps or system features, but they did not necessarily think that the systems were more secure and/or accurate than the paper ballot voting method. Further, the participants did not have comprehensive mental models that account for how the systems work. This contrasts with the complex conceptual models of security experts who are familiar with these systems.

*Study 3: To Verify or Not to Verify, That is The Question: Voters’ Anticipated Use and Expectations of Vote Verification Methods.* Many voters indicated that they want to verify their votes after participating in an election. With that being said, there was not a single form of verification that the participants preferred—about a quarter of the participants wanted a general confirmation that their votes were cast and counted, with roughly another fifty percent desiring more concrete information like a unique code
associated with each ballot selection, a list of their actual selections, and/or an image of their cast ballot. As the next generation of systems is developed, it will be a challenge to meet these diverse preferences without revealing individuals’ votes.

The findings in this research project are consistent with previous studies’ results. As described in the introduction (Chapter 1), other researchers have found significant usability issues with each system, and there were aspects of the systems that voters did not understand. This study greatly expanded upon this prior research by conducting comparative studies of usability using the ISO 9241-11 criteria and mental models across systems, thus allowing direct comparison of the results with both quantitative and qualitative data.

**Why aren't these systems more usable?**

It might be suggested that e2e systems with advanced cryptographic features are simply too complex for real voters to use. An interpretation like this would be misguided. While there are still a number of fundamental questions that require further research, I believe that significant improvements could be realized by implementing the simple changes suggested in Study 2.

There are four likely reasons that simple usability enhancing design improvements were overlooked: First, the focus of system design was on complex mechanics of central concern to security experts. When the research scientists developed the e2e voting systems, they likely focused on their expertise and ideas that they were concerned about—like system security and transparency. In contrast, there was not any evidence that they focused their efforts, with equal attention, on making the system
usable to the voter. Even if they did in fact make a substantial effort to ensure that the system is usable by voters, the system designers probably did not recognize that leaving out system usability as a key issue of concern, along with usability testing, throughout the entire design process contributed to the frustration voters experienced when using the systems.

Second, researchers have robust, conceptual models for the tested systems but failed to account for users who do not. Due to this misalignment between the system developers and users, real voters did not easily use the systems. Worded differently, “[d]esigners are not typical users” because they have close contact with the device and can operate it from their knowledgebase (Norman, 1988, p. 155). In contrast, users only have “knowledge in the world” (p. 157) so devices must be designed accordingly.

Third, there was little to no involvement of human factors engineers in the design cycle. People may think human factors are obvious or intuitive. However research has demonstrated that they are not obvious. They only seem obvious after principles or findings have been identified (Gould & Lewis, 1985). For this reason, it does not work to have an expert in a field like computer security try to apply human factors principles to a system that is being built, just like an engineering psychologist should not try to fix major security problems that a security expert could best and much more easily address. It is possible that the usability problems could have been entirely avoided simply by having a human factors expert work on the system alongside the security researchers.

Fourth, iterative system testing did not focus on the user. It is highly likely that the researchers who worked on the tested e2e systems used an iterative development method, but that iterative effort was consumed with the technical aspects of the security,
not the usability of the systems. An iterative process that included human factors would have likely identified misalignments between designers’ and users’ mental models that result in negative ramifications. Only by executing iterative development focused on the voters will the systems be efficient, effective, and satisfactory.

These four points explaining why the tested e2e voting systems were not usable are not new ideas. Human factors researchers have been calling attention to these types of problems for decades (e.g., Christensen, 1958), with the resulting mismatches having predictably bad consequences.

**Moving Forward With The Development of e2e Voting Systems**

To address these general human factors problems and move forward with developing an e2e system that is secure, transparent, auditable, reliable, and usable, an effort has been initiated in Texas called STAR-Vote (Bell et al, 2013). Using as much off the shelf computer equipment as possible, a system is being constructed to meet the needs of Travis County, Texas. The system is combining a user friendly DRE-style UI, printed paper ballots that are cast, and a voter-verifiable paper receipt with state-of-the-art cryptographic methods. To help ensure success, human factors / engineering psychologists, cryptographers, and auditing experts have been collaborating together. From the first day of development, the voting security and election auditing experts have been focusing on what they know best, while the human factors engineers are making certain that the systems are aligned with users’ cognitive and physical capabilities.

The STAR-Vote developers are also making the system appear as simple as possible to voters by hiding from all users as much of the complexity as possible—
resulting in a succinct and streamlined procedure that matches their previous voting experience. To vote cast and then vote verify with the STAR-Vote system, the voter follows these typical steps:

1. A voter checks in with a poll worker at the election-day vote center and is issued by the controller a 5-digit code linked to the type of ballot that they should be given, which is based on the voter’s precinct.

2. The voter approaches any available voting terminal and enters the code. They are then shown the ballot for their precinct.

3. The voter completes the ballot through a GUI or auditory UI (for non-sighted voters) that prevents typical voter errors like overvotes, undervotes, stray marks, and ambiguous selections. The voter also reviews his or her ballot on a final review screen for accuracy.

4. The system prints both a human-readable copy of the voter’s completed ballot for review and a receipt that includes information like the voting terminal used, time of vote, and a short code that serves as a link to the encrypted ballot.

5. The voter casts their ballot by first scanning in the serial number printed on the ballot and then dropping the ballot into the sealed box (some of the developers of STAR-Vote are discussing combining these two steps based on the findings of this research project). The voter keeps their receipt.

6. After the voter returns home, he or she may choose to navigate to the election’s public bulletin board and look for the posted ballot code. When
the voter sees it, the person knows their vote was transmitted by the system and will be, or has been, tallied to determine the election winner.

This is a simple procedure that matches closely as possible, voters’ current mental models composed of how-to-vote procedures. STAR-Vote’s developers are working diligently to strike a balance between hiding too much from the voter, resulting in a situation in which the system gives no indication of what it is doing or any type of feedback (Norman, 1988), and revealing too much information (the situation at present as demonstrated with the e2e systems tested in this project), resulting in user confusion or being overly burdened cognitively.

Even though the system is in its very early stages of development and has not been fully specified, the cross-disciplinary researchers are beginning to perform iterative testing that focuses simultaneously on improving cryptography, software engineering, usability, and auditing techniques. Usability experts want to make sure the system is effective, efficient, and satisfying for all voters, including those who have disabilities. The findings from these usability tests will be used to inform the next iteration of STAR-Vote. Research and design will be occurring in parallel (Kuniavsky, 2003; Nemeth, 2004), and this process of continual refinement through trial and error will allow researchers to meet the goals for the voting system in order to ultimately improve the “reliability, accuracy, fraud-resistance, and transparency of elections” (Bell et al, 2013, p. 35).
Future Research

Through this initial project, many opportunities for further research, independent of the development of STAR-Vote, have been identified. First, future work should determine if the modifications suggested in this research, particular those from part 2 of Study 1, actually mitigate the failures observed. Other usability issues should also be identified and addressed, such as vote casting and vote verification procedures that participants struggled with but ultimately overcame. Usability problems that disabled voters will encounter were also not studied. It is essential that all aspects of system interaction are satisfactory for all users. Even if a user does not fail at the task at hand, difficulty in use could deter voters from using it.

Other areas that need to be understood include how to develop systems that rely on common procedural knowledge, as voters should not be required to have full mental models to vote or verify their vote. Instead, they should be able to use the knowledge they already have about voting in general to be able to complete the vote casting process.

How users perceive trust and security should be understood so it can be leveraged to enhance users’ perceptions of these attributes. If voters find the systems to be secure and trustworthy, it might increase their confidence in the systems and subsequent system use. Optimizing system use is especially important in democratic elections in which every voice should be heard. Methods of trust and security enhancement might include security theater, or implementing features that make the system seem more trustworthy and secure without actually making it more secure. The use of invisible ink and decoder pens to reveal highlighted Scantegrity II codes is an example of security theater.
Participants expressed in their mental models interview this was a great security feature, when in fact printing the codes in invisible ink does nothing to increase actual system security. Or perhaps there might be better means, which have yet to be identified, to help users comprehend that the procedures they are executing are actually enhancing security. It is essential that if this is done that it does not add to the user’s cognitive workload or confuse them.

The best means to provide voters with verification instructions and convey information required to check on their votes needs to be identified. Many of the tested systems, particularly Helios and Scantegrity II, are both too cumbersome and too error prone to be fully effective as designed. For instance, Helios requires voters to either record by hand or print a long smart ballot tracker and then verify that this tracker matches the one showed later in the verification process. Comparing two long, intricate strings of numbers and case-sensitive characters requires a high amount of resources and is prone to human error. Scantegrity II is also problematic. The system requires voters to write down a code associated with every race selection and then compare these to the ones shown online. Even if Scantegrity II’s verification information was automatically generated for the users, it still would take a long time and imposes a substantial cognitive cost on the voter to compare all of the codes on the receipt to those shown on the verification website (this study only had 27 races on the ballot, but the Harris County, Texas 2010 November ballot had 88 races and three propositions—imagine recording a code by hand for each of these races and then comparing them to 91 codes found on a website!). As can be seen, the verification process needs to be streamlined as much as possible. Another problem is that the verification information displayed on the websites
needs to be helpful and meaningful. For instance, if voters are shown their cast selections (in the case of PaV) but cannot recognize deviations from their intentions, then the system should be revised. The ultimate aim when developing the verification systems should be to impose the least burden on the voter and minimize the opportunity for behavioral errors. At the same time, voters should be provided with the maximum amount of meaningful information. On top of all of this, the system must never reveal the voters’ selections.

Last, a method needs to be developed for voters to report identified system problems, along with protocols for election officials to follow. After voting in an election, a voter might verify his or her ballot on the election website, and in the process note that either their ballot is not posted or there is a discrepancy between the information shown on their receipt and that shown on the site. At this point, what are voters supposed to do to report these types of concerns? And what procedures do the election officials follow to determine if there are actual security problems, rather than a voter making an error like writing down the wrong ballot information? This aspect of the voting system has not been addressed in the published literature. Yet, all of the systems are designed to allow voters and non-voters to detect tampering and inaccuracies that can then be reported. Consequently, a fraud and error reporting system that is fully specified and tested needs to be developed. A reporting system, too, has the potential to have usability and human factors issues.
Conclusion

In summary, all of the voting security in the world does not matter if voters cannot use the voting system to cast their ballot. e2e voting systems must be usable along with being accurate, reliable, secure, and transparent. Only then can uncompromised, integral elections be realized. Human Factors researchers need to collaborate with security and auditing experts to realize the best systems possible for every user—regardless of variables like age, background, cognitive capability, and disability.
REFERENCES


Byrne, M. D., Greene, K. K., & Everett, S. P. (2007). Usability of voting systems: Baseline data for paper, punch cards, and lever machines. *Human Factors in


Help America Vote Act (HAVA) of 2002, Public Law 107-252.


Sisson, D. (n.d.). *Assumptions about user search behavior.* Retrieved from philosophe.com/search_topics/user_behavior/


APPENDIX A

Helios System and Study Materials

General Election
Harris County, Texas
November 8, 2016

To participate in this election, you will need to use the internet. For voting instructions, please go to: mail.google.com

Login to Gmail using the following information:

Username: xraychicken
Password: suitandtie

Figure A.1. Study Instructions for the Helios Mock-election

Figure A.2. Screenshot of the emailed instructions and link to the Helios Election
General Election, Harris County, Texas

To cast a vote, you will be led through the following steps.
If you have not yet logged in, you will be asked to do so at the very end of the process.

1. **Select** the answers you prefer.
   You can easily navigate forwards and backwards through the questions.

2. **Confirm** your selection.
   Your choices are encrypted safely inside your browser, and you get a smart ballot tracker.

3. **Submit** your encrypted ballot.
   Proceed to log in and cast your encrypted ballot for tallying.

Figure A.3. Screenshot of the Helios Voting Booth Instructions

General Election, Harris County, Texas

President
Question 1 of 27 — select up to 1 answer

- Gordon Bearer (R) (VP-Nathan Maclean)
- Vernon Stanley Albury (D) (VP-Richard Rigby)
- Janette Froman (L) (VP-Chris Aponte)

Maximum number of options selected.
To change your selection, please de-select a current selection first.

Figure A.4. Screenshot of the Presidential Race on the Helios Ballot
Question #25: Proposition 4: 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 hydropower, and hydrogen fuel cells; requiring that a percentage of retail electricity sales be derived from renewable sources, beginning with 3% in the year 2007 and increasing to 10% by 2015; requiring utilities to offer consumers a rebate of $2.00 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 retail 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 used to meet 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 this amendment?

[update]

Question #26: 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, 2007; 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?

[update]

Question #27: Proposition 6: 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?

[update]

Confirm Choices and Encrypt Ballot

Figure A.5. Screenshot of the Helios Review Screen

General Election, Harris County, Texas — Submit your Vote

We have received, but not yet recorded, your encrypted ballot.
Your smart ballot tracker is:

cIzxampJ0vDsiJB0/vZySlwU8Iv+GV2rdwZGDxw3i4k

Now, we need you to log in, so we can verify your eligibility.

[google]
[facebook]
[yahoo]

Don't worry, we'll remember your ballot while you log in.

Figure A.6. Screenshot of Helios Vote Received Page
General Election, Harris County, Texas — Vote Successfully Cast!

Congratulations, your vote has been successfully cast!

Your smart ballot tracker is:

clxampJ0vDsiJB0/v?ySlwU8Iv+GV2rdwZGDXw3i4k

For your safety, we have logged you out.

[ return to election info ]

**Figure A.7.** Screenshot of the Helios Cast Vote Confirmation Page

**Figure A.8.** Screenshot of Helios’ Voters and Ballot Tracking Center
Figure A.9. Screenshot of a Voter’s Archived Ballot (Accessed by Voter Through the Emailed Cast Ballot Confirmation Link)
APPENDIX B

PAV System and Study Materials

General Election Ballot
Harris County, Texas
November 8, 2016

INSTRUCTIONS TO VOTERS

1. Mark a cross (x) in the right hand box next to the name of the candidate you wish to vote for. For an example, see the completed sample ballot below. Use only the marking device provided or a number 2 pencil. Please note that this ballot has multiple cards. If you make a mistake, don’t hesitate to ask for a new ballot. If you erase or make other marks, your vote may not count.

2. After marking all of your selections, detach the candidates lists (left side of cards).

3. Shred the candidates lists.

4. Feed your voting slips into the scanner.

5. Take your receipts. Receipts can be used to confirm that you voted by visiting votingstudy.rice.edu.

Figure B.1. Voting Instructions for PAV
### Figure B.2. Card 1/8 of the PAV Ballot

<table>
<thead>
<tr>
<th>Card Key: 7rJ94K-1</th>
<th>Card 1 of 8</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>PRESIDENT AND VICE PRESIDENT</strong></td>
<td><strong>Ballot Continues on Card 2</strong></td>
</tr>
<tr>
<td><strong>President and Vice President</strong></td>
<td><strong>Vote for One</strong></td>
</tr>
</tbody>
</table>
| REP | Gordon Bearer  
Nathan Maclean |
| DEM | Vernon Stanley Albury  
Richard Rigby |
| LIB | Janette Froman  
Chris Aponte |

Mark a cross (X) in the right hand box next to the name of the candidate you wish to vote for.

<table>
<thead>
<tr>
<th><strong>CONGRESSIONAL</strong></th>
<th><strong>Vote for One</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>United States Senator</strong></td>
<td><strong>Vote for One</strong></td>
</tr>
<tr>
<td>REP</td>
<td>Cecile Cadieux</td>
</tr>
<tr>
<td>DEM</td>
<td>Fern Brzezinski</td>
</tr>
<tr>
<td>IND</td>
<td>Corey Dery</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Representative in Congress, District 7</strong></th>
<th><strong>Vote for One</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>REP</td>
<td>Pedro Brouse</td>
</tr>
<tr>
<td>DEM</td>
<td>Robert Mettler</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>STATE</strong></th>
<th><strong>Vote for One</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Governor</strong></td>
<td><strong>Vote for One</strong></td>
</tr>
<tr>
<td>REP</td>
<td>Glen Travis Lozier</td>
</tr>
<tr>
<td>DEM</td>
<td>Rick Stickles</td>
</tr>
<tr>
<td>IND</td>
<td>Maurice Humble</td>
</tr>
</tbody>
</table>
After polls close, you can check your votes online: votingstudy.rice.edu. Your ballot verification code is 7rJ94K.

**General Election Ballot**

**Harris County, Texas**

**November 8, 2016**

---

**Figure B.3.** PAV Voter Receipt

---
Texas General Election, November 8, 2016

Welcome!

This is a web page that can be used to check your votes that were cast in the November 2016 General Election in Harris County.

To verify your vote, please submit your ballot verification code here:

Submit

After all polls close, this page can be used to view the vote tally.

Figure B.4. Screenshot of PaV’s Vote Verification Home Webpage

Figure A2.4. Screenshot of PaV’s vote validation web page

Vote Validation Page

The verification code that you entered is valid.

Your vote has been recorded and registered. The images below show every selection on every card that makes up your ballot.

Once you have checked your vote, click here to to sign out.

Figure B.5. Screenshot of PaV’s Vote Validation Web Page
Appendix C

Scantegrity II System and Study Materials

Figure C.1. Front of Scantegrity II Ballot
Figure C.2. Photograph of a Completed Scantegrity II Ballot, with Invisible Ink Confirmation Codes Revealed
INSTRUCTIONS FOR VERIFYING YOUR VOTE ON-LINE AFTER YOU RETURN HOME

You have the **OPTION** of verifying your vote on-line after you return home. **It is not necessary to do so.** You may ignore this step entirely: **your cast ballot will be counted whether or not you do this verification process.**

If you wish to verify your vote on-line, perform the following steps:

1. Fill out your ballot according to the instructions provided on the ballot. “Confirmation numbers” will appear inside the oval you mark.

2. **BEFORE YOU CAST YOUR BALLOT** record the Online Verification Number and the confirmation numbers below, using the special pen.

   “On-Line Verification Number” from the bottom right corner of your ballot:

   ![On-Line Verification Number]

<table>
<thead>
<tr>
<th>Race</th>
<th>Code</th>
</tr>
</thead>
<tbody>
<tr>
<td>President And Vice President</td>
<td></td>
</tr>
<tr>
<td>United States Senator</td>
<td></td>
</tr>
<tr>
<td>Representative in Congress</td>
<td></td>
</tr>
<tr>
<td>Governor</td>
<td></td>
</tr>
<tr>
<td>Lieutenant Governor</td>
<td></td>
</tr>
<tr>
<td>Attorney General</td>
<td></td>
</tr>
<tr>
<td>Comptroller of Public Accounts</td>
<td></td>
</tr>
<tr>
<td>Commissioner of General Land Office</td>
<td></td>
</tr>
<tr>
<td>Commissioner of Agriculture</td>
<td></td>
</tr>
<tr>
<td>Railroad Commissioner</td>
<td></td>
</tr>
<tr>
<td>State Senator</td>
<td></td>
</tr>
<tr>
<td>State Representative District 134</td>
<td></td>
</tr>
<tr>
<td>Member State Board of Education, District 2</td>
<td></td>
</tr>
<tr>
<td>Judge Texas Supreme Court</td>
<td></td>
</tr>
<tr>
<td>Judge Court of Criminal Appeals</td>
<td></td>
</tr>
<tr>
<td>District Attorney</td>
<td></td>
</tr>
<tr>
<td>County Treasurer</td>
<td></td>
</tr>
<tr>
<td>Sheriff</td>
<td></td>
</tr>
<tr>
<td>County Tax Assessor</td>
<td></td>
</tr>
<tr>
<td>Justice of the Peace</td>
<td></td>
</tr>
<tr>
<td>County Judge</td>
<td></td>
</tr>
<tr>
<td>Proposition 1</td>
<td></td>
</tr>
<tr>
<td>Proposition 2</td>
<td></td>
</tr>
<tr>
<td>Proposition 3</td>
<td></td>
</tr>
<tr>
<td>Proposition 4</td>
<td></td>
</tr>
<tr>
<td>Proposition 5</td>
<td></td>
</tr>
<tr>
<td>Proposition 6</td>
<td></td>
</tr>
</tbody>
</table>

3. **Cast your ballot as usual using the polling station’s scanner. DO NOT CAST THIS SHEET, but take it home with you.**

4. After you have returned home, use a computer with an Internet connection to access the County’s vote verification web page: [mockelection.rice.edu](http://mockelection.rice.edu). Here you will see instructions for verifying that the confirmation numbers you wrote down are correctly recorded. Note that the confirmation numbers are randomly generated and cannot be used to determine how you voted.

**Figure C.3. Scantegrity II Vote Verification Sheet**
Figure C.4. Screenshot of Scantegrity II Vote Verification Home Page

Your ballot has been cast and processed! Below is the list of your confirmation numbers that correspond to your ballot’s selections.

Online Validation Code: HIC-2016-11-08-420795502

Confirmation Numbers for Cast Ballot:

- 4AB
- A2A
- 11K
- GL9
- Z31
- P6P
- JK3
- 713
- S1L
- 222
- FG3
- H11
- H15
- 1P9
- 499
- XRC
- SE6
- 1EK
- A31
- T7H
- 6C1
- 214
- 99L
- G9Y
- N9R
- 8LL
- 80Y

Click here when you are ready to log out of this page.

Figure C.5. Screenshot of Scantegrity II Cast Ballot Confirmation Page
Republican Directed Voting Slate

Please vote for the following candidates and propositions on each ballot.

President And Vice President:
Gordon Bearce and
Nathan Maclean (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:
Peta 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 Directed Voting Slate

Please vote for the following candidates and propositions on each ballot.

<table>
<thead>
<tr>
<th>Position</th>
<th>Candidate</th>
</tr>
</thead>
<tbody>
<tr>
<td>President And Vice President</td>
<td>Vernon Stanley Albury and Richard Rigby (D)</td>
</tr>
<tr>
<td>United States Senator:</td>
<td>Fern Brzezinski (D)</td>
</tr>
<tr>
<td>Representative in Congress:</td>
<td>Robert Mettler (D)</td>
</tr>
<tr>
<td>Governor:</td>
<td>Rick Stickles (D)</td>
</tr>
<tr>
<td>Lieutenant Governor:</td>
<td>Shane Terrio (R)</td>
</tr>
<tr>
<td>Attorney General:</td>
<td>Rick Organ (D)</td>
</tr>
<tr>
<td>Comptroller of Public Accounts:</td>
<td>Greg Converse (D)</td>
</tr>
<tr>
<td>Commissioner of General Land</td>
<td>Sam Saddler (R)</td>
</tr>
<tr>
<td>Commissioner of Agriculture:</td>
<td>Roberto Aron (D)</td>
</tr>
<tr>
<td>Railroad Commissioner:</td>
<td>Zachary Minick (D)</td>
</tr>
<tr>
<td>State Senator:</td>
<td>Wesley Steven Millette (D)</td>
</tr>
<tr>
<td>State Representative District 134</td>
<td>Susanne Rael (D)</td>
</tr>
<tr>
<td>Member State Board of Education</td>
<td>Mark Baber (D)</td>
</tr>
<tr>
<td>District 2:</td>
<td></td>
</tr>
<tr>
<td>Presiding Judge Texas Supreme</td>
<td>Tim Graasy (D)</td>
</tr>
<tr>
<td>Court Place 3:</td>
<td></td>
</tr>
<tr>
<td>Presiding Judge Court of Criminal Appeals Place 2:</td>
<td>Dan Plouffe (R)</td>
</tr>
<tr>
<td>District Attorney:</td>
<td>Jennifer A. Lundeer (D)</td>
</tr>
<tr>
<td>County Treasurer:</td>
<td>Gordon Kallas (D)</td>
</tr>
<tr>
<td>Sheriff:</td>
<td>Jason Valle (L)</td>
</tr>
<tr>
<td>County Tax Assessor:</td>
<td>Randy H. Clemons (CON)</td>
</tr>
<tr>
<td>Justice of the Peace:</td>
<td>Clyde Gayton Jr.</td>
</tr>
<tr>
<td>County Judge:</td>
<td>Lewis Shine</td>
</tr>
<tr>
<td>Proposition 1:</td>
<td>Yes</td>
</tr>
<tr>
<td>Proposition 2:</td>
<td>No</td>
</tr>
<tr>
<td>Proposition 3:</td>
<td>Yes</td>
</tr>
<tr>
<td>Proposition 4:</td>
<td>No</td>
</tr>
<tr>
<td>Proposition 5:</td>
<td>Yes</td>
</tr>
<tr>
<td>Proposition 6:</td>
<td>Yes</td>
</tr>
</tbody>
</table>
APPENDIX E

System Usability Scale (SUS) Adapted for Voting

Please check the box that reflects your immediate response to each statement. Don’t think too long about each statement. Make sure you respond to every statement. If you don’t know how to respond, simply check box “3.”

<table>
<thead>
<tr>
<th>Statement</th>
<th>Strongly Disagree</th>
<th>Strongly Agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. I think that I would like to use this voting system frequently.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. I found the voting method unnecessarily complex.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. I thought the ballot was easy to use.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. I think that I would need the support of a poll worker to be able to use this system.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. I found the various parts of this voting method were well integrated.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. I thought there was too much inconsistency in this voting system.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. I would imagine that most people would learn to use this voting system very quickly.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. I found the voting method very awkward to use.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9. I felt very confident using the system.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10. I needed to learn a lot of things before I could get going with this voting method.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
## SUS Adapted for Vote Verification

Please check the box that reflects your immediate response to each statement. Don’t think too long about each statement. Make sure you respond to every statement. If you don’t know how to respond, simply check box “3.”

1. I think that I would like to use this verification system frequently.

<table>
<thead>
<tr>
<th>Strongly Disagree</th>
<th>Strongly Agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>5</td>
<td></td>
</tr>
</tbody>
</table>

2. I found the verification system unnecessarily complex.

   | 1                 | 2             |
   | 3                 | 4             |
   | 5                 |               |

3. I thought the verification system was easy to use.

   | 1                 | 2             |
   | 3                 | 4             |
   | 5                 |               |

4. I think that I would need the support of a technical person or poll official to be able to use this verification system.

   | 1                 | 2             |
   | 3                 | 4             |
   | 5                 |               |

5. I found the various parts of this system were well integrated.

   | 1                 | 2             |
   | 3                 | 4             |
   | 5                 |               |

6. I thought there was too much inconsistency in this verification system.

   | 1                 | 2             |
   | 3                 | 4             |
   | 5                 |               |

7. I would imagine that most people would learn to use this verification system very quickly.

   | 1                 | 2             |
   | 3                 | 4             |
   | 5                 |               |

8. I found the verification system very awkward to use.

   | 1                 | 2             |
   | 3                 | 4             |
   | 5                 |               |

9. I felt very confident using the verification system.

   | 1                 | 2             |
   | 3                 | 4             |
   | 5                 |               |

10. I needed to learn a lot of things before I could get going with this verification system.

   | 1                 | 2             |
   | 3                 | 4             |
   | 5                 |               |
APPENDIX F

Final Participant Survey

Note: Survey was administered for studies 1, 2, and 3

Voting Study Survey Packet

2013

Participant Number

Date
Background Information

1. **Age:** _______ (years)

2. **Gender:** _____ Male _____ Female

3. **Do you have normal or corrected to normal vision?** ____ No ____ Yes

4. **Do you consider yourself to have a reading disability?** ____ No ____ Yes

5. **Are you left or right handed?** ____ Right ____ Left ____ Ambidextrous

6. **Are you a native English speaker?** ____ No ____ Yes

   If no, what is your native language? __________________________

7. **What is your current occupation?** _________________________

8. **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.)
9. 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: ____________________________

10. Which of the following income ranges best describes your yearly wages?
   ___ below $20,000
   ___ $20,000 to $40,000
   ___ $40,000 to $60,000
   ___ $60,000 to $80,000
   ___ Above $80,000
11. If you are retired, which of the following income ranges best describes your maximum yearly wages while you were working full-time?

___ below $20,000
___ $20,000 to $40,000
___ $40,000 to $60,000
___ $60,000 to $80,000
___ Above $80,000
___ I am not retired

12. Can you touch type? (Can you type without looking at the keys?)

___ No ___ Yes

13. 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

14. Please rate your level of computer expertise.

Circle your answer, with 1 = novice and 10 = expert.

1 2 3 4 5 6 7 8 9 10
15. **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)
- [ ] Other—please specify: _________________________________
16. 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)
Previous Voting Information

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

18. How many national elections have you voted in? ________

19. In which state(s) and county(s) have you voted in a national election?
   _______________________________________________________
   _______________________________________________________

20. How many other elections of any type (local, school board, etc.) have you voted in? ________

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

***For questions 22 - 29, please keep in mind your previous voting experience in any type of election (not including the voting you did in this study). If you have never voted, please skip the remaining questions in this section and proceed to page 10.***

22. **Do you typically cast your vote on an absentee ballot?**
   
   _____ No _____ Yes

23. **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
   _____ Punchcards
   _____ Electronic – touchscreen
   _____ Electronic – other
   _____ Don’t know
   _____ Other—please specify:__________________________
24. Have you ever felt worried about figuring out how to use the ballot or other technology to cast your vote? ____ No ____ Yes

25. 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

26. Did this prevent you from voting? ____ No ____ Yes

27. Do you typically vote a straight-party ticket (in other words, do you typically vote for candidates from only one party)? ____ No ____ Yes

28. Do you typically cast a vote for every office on the ballot? ____ No ____ Yes

29. 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.

______________________________
______________________________
______________________________
______________________________
______________________________

**** End of previous voting experience section****
30. 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

31. If you have been following the news about computer voting and security, has it affected your trust of these systems?
   ___ No
   ___ Yes
   ___ Does not apply to me

   Why or why not? __________________________________________________________
   __________________________________________________________
   __________________________________________________________
   __________________________________________________________
   __________________________________________________________
   __________________________________________________________
   __________________________________________________________
Voting Method Comparison Survey

32. Of the three voting methods you used in this study, which was your favorite? Please circle one.

Invisible Ink      Computer      Cards w/ detachable lists

Why was this your favorite voting system? ________________________
________________________________________________________
________________________________________________________

33. Of the three voting methods you used in this study, which was your least favorite? Please circle one.

Invisible Ink      Computer      Cards w/ detachable lists

Why was this your least favorite type? ________________________
________________________________________________________
________________________________________________________

34. What, if anything, did you not like about the least favorite voting method (the kind that was not your favorite)? ______________
________________________________________________________
________________________________________________________
35. **Have you ever voted on another type of ballot or voting system that you liked better than these?**  
   ____ No  ____ Yes  
   If yes, please describe the other ballot. ____________________________  
   ____________________________  
   ____________________________  

36. **Circle the voting methods that you would want to use to vote in the future.**  
   Invisible Ink  Computer  Cards w/ detachable lists  
   e-Slate  Other—please specify: ____________________________
Security Comparison Survey

37. Rank the security of the following systems, with 1 = the most secure system and 3 = the least secure system.

[ ] Invisible Ink
[ ] Computer
[ ] Cards with Detachable Lists

38. Please respond to the following statement: I am confident that the invisible ink voting system is secure.

Circle your answer, with 1 = not secure at all and 10 = very secure.

1 2 3 4 5 6 7 8 9 10

I feel the invisible ink voting method is secure, because 

________________________________________________________________________

________________________________________________________________________

I feel the invisible ink voting method is unsecure, because 

________________________________________________________________________

________________________________________________________________________.
39. Please respond to the following statement: I am confident that the computer voting system is secure.

Circle your answer, with 1 = not secure at all and 10 = very secure.

1  2  3  4  5  6  7  8  9  10

I feel computer voting method is secure, because __________________________

______________________________

______________________________.

I feel the computer voting method is unsecure, because __________

______________________________

______________________________.

40. Please respond to the following statement: I am confident that the cards with detachable lists voting system is secure.

Circle your answer, with 1 = not secure at all and 10 = very secure.

1  2  3  4  5  6  7  8  9  10

I feel the cards with detachable lists voting method is secure, because __________________________

______________________________

______________________________.

I feel the cards with detachable lists voting method is unsecure, because __________________________
41. **Have you ever voted on another type of ballot or voting equipment that you felt was more secure than these?**  
   ____ No  ____ Yes  

   If yes, please describe the other ballot or voting system.

   __________________________________________________________________________

   __________________________________________________________________________

   __________________________________________________________________________

   Specifically, what made this ballot or voting system more secure?

   __________________________________________________________________________

   __________________________________________________________________________

   __________________________________________________________________________
Instructions Survey

42. Did you feel that the instructions for the invisible ink voting method were easy to understand? ___ No ___ Yes
   
   If no, what was unclear?

43. Did you feel that the instructions for the computer voting method were easy to understand? ___ No ___ Yes

   If no, what was unclear?

44. Did you feel that the instructions for the cards with detachable lists voting method were easy to understand? ___ No ___ Yes

   If no, what was unclear?
Vote Verification Survey

45. Why do the voting methods that you used today give you the option to check your votes after your vote has been cast?

_______________________________________
_______________________________________
_______________________________________
_______________________________________

46. After voting, do you think it is important to be able to verify that your vote was cast and counted as intended?

Circle your answer, with 1 = not at all important and 10 = very important.

1  2  3  4  5  6  7  8  9  10

Why or why not? __________________________
_______________________________________
_______________________________________

47. If this was a real election and you could go home now and verify on the internet that your vote was cast and counted as you intended, how likely is it that you would do this?

Circle your answer, with 1 = not at all likely and 10 = very likely.

1  2  3  4  5  6  7  8  9  10

OR I don’t have internet
48. Which of the following would you want to see on a website that allows you to verify your vote? (Check all that apply)

____ A digital image of your completed ballot
____ A list of the names of all the people you voted for
____ Confirmation that your ballot was cast and counted in the total
____ A confidential PIN number that is associated with each candidate you voted for, and you are the only person who knows what the PIN number means
____ None
____ Don’t know

49. Place a check next to the systems that you would want to check votes for if this was a real election.

____ Invisible Ink
____ Computer
____ Cards with detachable lists