Users' Mental Models for Three End-to-End Voting Systems: Helios, Prêt à Voter, and Scantegrity II

Claudia Z. Acemyan[1(✉)], Philip Kortum[1], Michael D. Byrne[1,2], and Dan S. Wallach[2]

[1] Department of Psychology, Rice University, Houston, TX, USA
{claudiaz,pkortum}@rice.edu
[2] Department of Computer Science, Rice University, Houston, TX, USA
{byrne,dwallach}@rice.edu

Abstract. This study sought to understand voter's mental models for three end-to-end (e2e) voting systems: Helios, Prêt à Voter, and Scantegrity II. To study voters' mental models of e2e systems, 16 Houston area voters participated in mock elections that required them to vote first with a paper ballot and then with the three e2e systems. After using each system, subjects were asked to draw their mental model—or how the system works, then describe it to the experimenter, and last complete an interview. We found that most participants think about the systems first and foremost in terms of how-to-vote procedures, rather than detailed, conceptual models that describe all aspects of a system, including how they work. When designing e2e voting systems, the findings from this study can be used by system developers to ensure that voters find the systems easy to use and that the designs align with voters' pre-existing mental models for voting.

Keywords: Mental models · Voting systems · End-to-end voting methods · User centered design

1 Introduction

Problems often arise when system developers expect future system users to have a specific mental model for a system when in fact they will have another [1]. Differences between the two models may lead to users making errors while trying to complete a task, becoming frustrated, taking longer to complete goals, failing to recognize why they have to execute procedures, using the system differently than conceptualized, and avoiding using the system all together despite its advantages. Even with this knowledge, system developers continue to make assumptions—from how future users will think about the system to how they will use it—rather than using empirical data that can be used to optimize principles like usability, learn ability, satisfaction, expectations, and problem solving.

An example of a system type in which many of the developers made assumptions about the users' mental models is end-to-end (e2e) voting systems. E2e voting systems were designed to improve the integrity of elections by ensuring that voters can cast

T. Tryfonas and I. Askoxylakis (Eds.): HAS 2015, LNCS 9190, pp. 463–474, 2015.
DOI: 10.1007/978-3-319-20376-8_41

intended vote selections using a reliable, cryptographically secure system and that these votes will then be counted accurately as cast. The systems also provide a means to audit the system by election officials, voters, and any interested third party. For instance, if a voter wants to make sure their ballot is recorded by the system, many e2e systems give them a way to do this through a website after polls close.

In the process of developing e2e voting systems, voting experts also wanted to enhance voters' awareness of system security, make how the systems work more transparent, increase accuracy by having voters be active participants in the voting process, 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 theoretical frameworks that set 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" [2]—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, and they also believed it would be important to others, especially voters. Accordingly, the Helios team worked to realize a system that voters could trust because they thought it would be 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" [3], 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" [4]. To achieve this end, the system has features that allow the voters to check that it is accurate and performing as expected. For example, when a voter makes a ballot selection with a special decoder pen, 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 e2e voting methods 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 (e.g., casting a ballot by placing it in a ballot box versus scanning it), or perhaps voters might not know what they are

doing and/or why (e.g., voters might not understand why PaV requires them to essentially tear their ballot in half, shred the side with candidate names on it, and then scan the half with their selections).

These issues are of concern not only to human factors researchers, but also to voting security experts. Ben Adida 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 [2]. Moreover, Adida hypothesized that a system like Helios is not widely recognized by its users to represent a major improvement over methods currently used in elections. The means 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 that users form after vote casting and vote verifying with the Helios, Prêt à Voter, and Scantegrity II methods. These systems were selected because they are representative of the different types of e2e systems, are widely accepted as viable options by the voting research community, and have been used in elections [5]. Through this research, there was also a desire to determine if these models support the security concepts integral to the e2e systems. Glimpses 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 actions that are required to cast a ballot or check on it after the election? And do voters find the systems to be more secure, trustworthy, and 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 [6]), the exclusive focus of this body of work was on assessing system usability, not on understanding user mental models.

In contrast, there is an extensive, diverse body of work on mental models. A single definition for mental models does not exist, as the term seems to be continually redefined across researchers and projects. Hence, in this paper, a mental model for a voting system is defined using a combination of the previously published definitions [7–9]: 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.

2 Research Study

2.1 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. 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: 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.

So that voters knew for whom they should vote, they were randomly assigned one of two lists of candidates and propositions. Their list was either primarily Republican or it was primarily Democratic. These two lists were the same as those used in previous voting studies conducted at Rice University [5, 10].

The main dependent variable was the mental model formed for each voting system. 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. In addition, basic demographic and background information was 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. 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 of system 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 the paper ballot voting method?" This question was asked to determine if the participants thought that the system was more secure than a non-e2e voting method. 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.

Materials. The following is a general summary of study materials. For detailed information, refer to our previous paper [5]. The ballot was composed of 21 races and six propositions. These were the same candidates and propositions used in Rice University's previous voting studies. The Helios voting system ran through Helios' website at https://vote.heliosvoting.org during the summer of 2013. The PaV system was developed by consulting published papers and its website [3, 11–13] and through discussions with Peter Ryan, one of the primary developers. The Rice version of Scantegrity II was based on 2009 Takoma Park printed election materials [e.g., 16, 17], published articles about the system [e.g., 3, 16, 18], and in consultation with both Aleks Essex and Richard Carback who were involved with the development and implementation of the system (A. Essex, personal communication, December 14, 2012; R. Carback, personal communication, December 13, 2012).

2.2 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 Fig. 1 for a representative example. As for the subjects who did not focus on the how-to-vote procedures, one participant drew a different symbol for each of the voting systems that expressed their impressions of the system. Another participant 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 [8], 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 the drawn models were inaccurate at times (e.g., some voters thought Helios included steps to make sure that they were a human voter, versus a computer system).

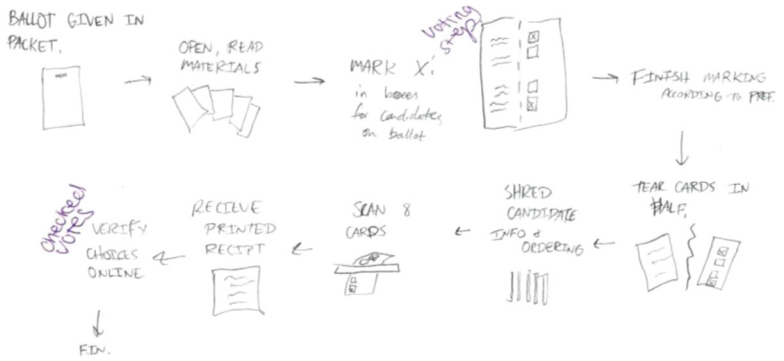

Fig. 1. One example of a drawn mental model for PaV

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. 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.

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. The listed secure features 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, few participants (6 %) indicated this level of understanding. 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, 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. 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.

Eighty-eight percent 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.

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 despite a long, block of instructions presented to voters at the top of the ballot. 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.

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.

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.

To summarize the security findings, voters generally 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.

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. Many of the cited components across systems 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 different conceptual model for the e2e systems from the actual voters.

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, of 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. 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 procedures, with the minority dealing with the reason why a step was required to vote or check on their vote.

System Preferences. When asked which of the four methods participants would prefer to use in a real election and why, 38 % preferred Helios. 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, χ^2 (3, N = 16) = 3.5, $p = 0.32$. When asked which of the four methods participants would least prefer to use in a real election, participants least preferred Scantegrity II, with 38 %. 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 felt that overall the system was confusing. Again, there was no evidence to support a difference in proportions of responses across systems, χ^2 (3, N = 16) = 3.5, $p = 0.32$. 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.

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. 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.

3 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. In addition, participants' mental models do not significantly change depending on the voting method used and at times are inaccurate.

Perhaps the static nature of e2e mental models is the product of voters relying on a global model for voting in a democratic election, which is then enhanced by the *specific voter's* rote procedures of voting with a specific method. Instead of developing

a mental model for how each system works, voters appeared to refer to the larger construct of the voter's voice being heard by going to a polling station to fill out an anonymous ballot, placing it in a secure ballot box, having the votes counted carefully by honest election officials, and then this tally determining the election outcome. They did not describe how the system worked, various system states and responses to user actions, or any of the theories and concepts detailed by the developers in their papers. Embedded within the general voting framework was the "how to cast a ballot" steps for the system at hand. If this theory holds true in future studies, it might also account for inaccuracies expressed in the mental models, such as leaving entire steps out, which make the mental model more generic. It would also explain why voters made errors like trying to cast a vote with PaV and Scantegrity II by placing a ballot in the ballot box, instead of first scanning it [5].

The nature of the mental models found in this study are also likely a product of many participants' only exposure to these e2e systems being this study. The subjects read lengthy instructions telling 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 systems. 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 as they were not explicitly apparent during system use.

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 either vote by mail or use the voting method at their designated polling station). As such, voters in real elections would encounter the same predicament as the participants in this test, and they 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, helps 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 used to keep their votes completely private yet trackable, then perhaps voters will recognize their value. In turn, these enhanced mental models might then lead to higher usability scores and better user performance on these systems.

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 and Bovair [17]). Similarly, voters just need to be able to get through the how-to-vote sequence 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 authors' conceptualization of these participants' mental models [8]. 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, yet there is still utility in studying them [17].

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 of U.S. voters since half of the subjects were Rice University students. Nevertheless, even the students are real voters, and the collected data offer a glimpse into how people might think about the voting systems.

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, some unique system-specific features, and principles the general democratic voting process. 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 e2e voting methods. This research should also serve as a friendly reminder to all system developers to collect data from actual system users, rather than make assumptions about how those users will think about and use the system.

Acknowledgements. This research has been supported in part by NSF award CNS-1049723. We would also like to thank Molly Ahn for assisting with the study.

References

1. Nielsen, J.: Mental Models. http://www.nngroup.com/articles/mental-models/
2. Adida, B.: Helios: Web-based open-audit voting. In: USENIX Security Symposium, pp. 335–348 (2008)
3. Ryan, P.Y.A., Bismark, D., Heather, J., Schneider, S., Xia, Z.: Prêt à voter: a voter-verifiable voting system. IEEE Trans. Inf. Forensics Secur. **4**, 662–673 (2009)
4. Chaum, D., Essex, A., Carback, R., Clark, J., Popoveniuc, S., Sherman, A., Vora, P.: Scantegrity: End-to-end voter-verifiable optical-scan voting. IEEE Secur. Priv. **6**, 40–46 (2008)

5. Acemyan, C.Z., Kortum, P., Byrne, M.D., Wallach, D.S.: Usability of voter verifiable, end-to-end voting systems: baseline data for Helios, Prêt à Voter, and Scantegrity II. USENIX J. Elect. Technol. Syst. **2**, 26–56 (2014)
6. Winckler, M., Bernhaupt, R., Palanque, P., Lundin, D., Leach, K., Ryan, P., Alberdi, E., Strigini, L.: Assessing the usability of open verifiable e-voting systems: a trial with the system Prêt à voter. In: Proceedings of ICEGOV 2009. pp. 281–296 (2009)
7. Rouse, W.B., Morris, N.M.: On looking into the black box: Prospects and limits in the search for mental models. Psychol. Bull. **100**, 349–363 (1986)
8. Norman, D.A.: Some observationson mental models. In: Gentner, D., Stevens, A.L. (eds.) Mental Models, pp. 7–14. Lawrence Erlbaum Associates, Hillsdale (1983)
9. Rasmussen, J.: On the structure of knowledge: a morphology of mental models in a man-machine system context. Risø National Laboratory, Risø (1979)
10. Byrne, M., Greene, K., Everett, S.: Usability of voting systems: Baseline data for paper, punch cards, and lever machines. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 171–180, ACM (2007)
11. Lundin, D., Ryan, P.Y.: Human readable paper verification of Prêt à voter. In: Jajodia, S., Lopez, J. (eds.) ESORICS 2008. LNCS, vol. 5283, pp. 379–395. Springer, Heidelberg (2008)
12. Ryan, P.Y., Peacock, T.: A threat analysis of Prêt à voter. In: Chaum, D., Jakobsson, M., Rivest, R.L., Ryan, P.Y., Benaloh, J., Kutylowski, M., Adida, B. (eds.) Towards Trustworthy Elections, New Directions in Electronic Voting. LNCS, vol. 6000, pp. 200–215. Springer, Heidelberg (2010)
13. Ryan, P.Y., Schneider, S.A.: Prêt à voter with re-encryption mixes. In: Gollmann, D., Meier, J., Sabelfeld, A. (eds.) ESORICS 2006. LNCS, vol. 4189, pp. 313–326. Springer, Heidelberg (2006)
14. Carback, R., Chaum, D., Clark, J., Conway, J., Essex, A., Herrnson, P., Mayberry, T., Popoveniuc, S., Rivest, R.L., Shen, E., Sherman, A., Vora, P.L.: Scantegrity II municipal election at Takoma Park: the first E2E binding governmental election with ballot privacy. In: Proceedings of the 19th USENIX Security Symposium (2010)
15. The City of Takoma Park Maryland. http://www.takomaparkmd.gov
16. Sherman, A.T., Carback, R., Chaum, D., Clark, J., Essex, A., Herrnson, P.S., Mayberry, T., Popoveniuc, S., Rivest, R.L., Shen, E., Sinha, B., Vora, P.: Scantegrity mock election at Takoma Park. In: Proceedings of the NIST Workshop on End-to-end Voting Systems, pp. 45–61 (2009)
17. Kieras, D.E., Bovair, S.: The role of a mental model in learning to operate a device. Cogn. Sci. **8**, 255–273 (1984)